

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 ; Search time 6250.78 Seconds  
(Without alignments)  
3295.534 Million cell updates/sec

Title: US-09-806-955-3

Perfect score: 1917

Sequence: 1 atggagagagagacaagaaga.....tcaccaccaccaccaccac 1917

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
22703874

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estda:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666.8	86.9	2539	12 AK004578	AK004578 Mus muscu
2	860.2	44.9	861	10 AL552861	AL552861 Mus muscu
3	812	42.4	825	11 BG675308	BG675308 602621612
4	812	42.4	1100	10 AL541667	AL541667 AL541667
5	796.8	41.6	814	11 BG681177	BG681177 602629315
6	770.6	40.2	777	11 BG743404	BG743404 602634914
7	754	39.3	782	11 B1260800	B1260800 602970779
8	747.4	39.0	946	10 BE740540	BE740540 601595632
9	735	38.3	736	11 BG674623	BG674623 602620671
10	731	38.1	769	10 BE742027	BE742027 601594538
11	726.6	37.9	1055	11 BE876797	BE876797 60148373
12	724.6	37.8	768	10 A1909687	A1909687 PM-BT219-

13	721.6	37.6	985	11 B1086390	B1086390 602849835
14	719.8	37.5	835	11 BG698078	BG698078 602659944
15	718.6	37.5	813	11 BG679349	BG679349 602628541
16	716.8	37.4	1205	11 BF792328	BF792328 602252860
17	711	37.1	967	10 AL555722	AL555722 AL555722
18	706.6	36.9	1003	10 AL569021	AL569021 AL569021
19	703.6	36.7	961	11 BG110618	BG110618 602281078
20	702	36.6	976	10 AL546661	AL546661 AL546661
21	681.4	35.5	874	10 BE740407	BE740407 601595436
22	672.8	35.1	775	11 BG118890	BG118890 602347893
23	668.4	34.9	778	11 BF341614	BF341614 602016039
24	667.6	34.8	838	10 BE746886	BE746886 601579381
25	665.6	34.7	918	10 AL558058	AL558058 AL558058
26	662.4	34.6	935	11 BG116226	BG116226 602318446
27	661	34.5	827	11 BG253599	BG253599 602364334
28	660.2	34.4	762	11 BG874601	BG874601 RC3-BN005
29	659.8	34.4	912	11 BF984440	BF984440 602307982
30	659.4	34.4	706	10 AU136376	AU136376 AU136376
31	659.4	34.4	829	11 BG684785	BG684785 602799210
32	659.2	34.4	919	11 BF311360	BF311360 601896761
33	656.2	34.2	923	11 BG865362	BG865362 602783969
34	655.2	34.2	807	11 B1101144	B1101144 602886829
35	654.8	34.2	829	11 BG923286	BG923286 602825247
36	649.8	33.9	766	11 BG681518	BG681518 602795482
37	649.8	33.9	923	11 BG255899	BG255899 602367542
38	648.6	33.8	857	11 B1218638	B1218638 602937865
39	647.6	33.8	924	10 AL542474	AL542474 AL542474
40	647	33.8	953	11 BG289393	BG289393 602381421
41	646.2	33.7	1019	10 AL578162	AL578162 AL578162
42	642.6	33.5	682	11 BG115553	BG115553 602317239
43	642	33.5	673	11 BG283834	BG283834 602407558
44	640.4	33.4	763	10 AV755667	AV755667 AV755667
45	639.6	33.4	893	10 BE790126	BE790126 601482941

#### ALIGNMENTS

RESULT 1  
AK004578  
LOCUS  
DEFINITION  
AK004578  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
3 (bases 1 to 2539)  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, D.,



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 ; Search time 150.12 seconds  
(without alignments)  
2892.075 Million cell updates/sec

Title: US-09-806-955-3

Perfect score: 1917  
Sequence: 1 atggagagagacagaagga.....ttccaccaccaccaccacac 1917

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11328999 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PTCUTS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	783.8	40.9	2574	1 US-08-441-139-6	Sequence 6, Appl
2	754	39.3	2780	1 US-08-441-139-1	Sequence 1, Appl
3	719	37.5	2089	1 US-08-441-139-13	Sequence 13, Appl
4	712	37.1	2379	4 US-08-797-358B-2	Sequence 2, Appl
5	638.6	33.3	2394	1 US-08-797-358B-1	Sequence 1, Appl
6	636.6	33.2	3607	1 US-08-647-351B-1	Sequence 1, Appl
7	548.4	28.6	5470	1 US-08-441-139-12	Sequence 12, Appl
8	548.4	28.6	5470	6 5196523-5	Patent No. 5196523
9	469.4	24.5	2155	1 US-08-214-563-1	Sequence 1, Appl
10	429	22.4	2515	4 US-09-066-047-4	Sequence 4, Appl
11	410	21.4	2367	1 US-08-441-139-3	Sequence 3, Appl
12	398.2	20.8	4320	2 US-08-472-534-4	Sequence 4, Appl
13	393.8	20.5	4411529	4 US-09-103-840A-1	Sequence 1, Appl
14	392.4	20.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl
15	350.6	18.3	931	1 US-08-203-905B-4	Sequence 4, Appl
16	326.6	17.0	1800	6 5240706-2	Patent No. 5240706
17	279.6	14.6	966	1 US-08-257-073-6	Sequence 6, Appl
18	238	12.4	5408	1 US-08-441-139-15	Sequence 15, Appl
19	200.8	10.5	618	4 US-09-328-111-233	Sequence 233, App
20	196.8	10.3	557	4 US-09-328-111-188	Sequence 198, App
21	189.4	9.9	681	4 US-09-328-111-288	Sequence 288, App
22	177.4	9.3	615	4 US-09-328-111-200	Sequence 200, App
23	171.8	9.0	1021	3 US-08-714-918-70	Sequence 70, Appl
24	171.8	9.0	1021	3 US-08-265-315-70	Sequence 70, Appl
25	171.8	9.0	1021	4 US-09-265-315-70	Sequence 70, Appl
26	171.8	9.0	1021	4 US-09-266-417-70	Sequence 70, Appl
27	144	7.5	439	1 US-08-644-729-19	Sequence 19, Appl

28	136.6	7.1	439	1 US-08-644-729-27	Sequence 27, Appl
29	135	7.0	439	1 US-08-644-729-15	Sequence 15, Appl
30	134.4	7.0	439	1 US-08-644-729-20	Sequence 20, Appl
31	133.8	7.0	439	1 US-08-644-729-17	Sequence 17, Appl
32	131.8	6.9	439	1 US-08-644-729-16	Sequence 16, Appl
33	130.2	6.8	439	1 US-08-644-729-21	Sequence 21, Appl
34	128.8	6.7	2245	1 US-08-203-905B-1	Sequence 1, Appl
35	127	6.6	439	1 US-08-644-729-25	Sequence 25, Appl
36	125.4	6.5	1731	1 US-08-203-905B-13	Sequence 13, Appl
37	125	6.5	439	1 US-08-644-729-22	Sequence 22, Appl
38	122.8	6.4	242	1 US-08-464-164-3	Sequence 3, Appl
39	122.8	6.4	242	1 US-08-338-057-3	Sequence 3, Appl
40	122.8	6.4	242	2 US-08-668-416-3	Sequence 3, Appl
41	121	6.3	439	1 US-08-644-729-28	Sequence 28, Appl
42	119.8	6.2	792	1 US-08-203-905B-3	Sequence 13, Appl
43	108.8	5.7	349	1 US-08-275-225-9	Sequence 9, Appl
44	105.6	5.5	250	1 US-08-464-164-5	Sequence 5, Appl
45	105.6	5.5	250	1 US-08-338-057-5	Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-08-441-139-6  
Sequence 6, Application US/08441139  
Patent No. 5773245  
GENERAL INFORMATION:  
APPLICANT: Wiltup, Dr. Karl D.  
APPLICANT: Robinson, Anne S.  
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF  
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,139  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/089,997  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8646  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SAMS UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2574 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 441..2429  
US-08-441-139-6





XX MPI: 2000-317942/27.  
DR P-PSDB: AAY90693.

XX New human immunoglobulin heavy chain binding protein and encoding  
PT polynucleotide, useful for diagnosis and treatment of rheumatoid  
PT arthritis

PS Claim 12: Page 24; 53pp; English.

XX This sequence represents cDNA encoding a human immunoglobulin heavy chain  
CC binding protein, BIP(78K), having a 639 amino acid sequence (Y90693).  
CC The invention also encompasses a BIP(78K) protein of 633 amino acids  
CC (Y90694). The cDNA encoding BIP(78K), also referred to as p78 in the  
CC specification, was isolated from human chondrocytes (the specialised  
CC cells of articular cartilage) and human chondrosarcoma cell lines. The  
CC BIP(78K) cDNA of this invention contains a number of differences  
CC compared with the published sequence (Genbank accession number X87949),  
CC and has therefore been deposited with Genbank with the accession  
CC number AF186111. These differences comprise 6 single nucleotide  
CC substitutions and an arginine insertion at position 834-836 of the  
CC protein. The BIP(78K) proteins react with antibodies present in the  
CC serum of rheumatoid arthritis patients, and is therefore a putative  
CC autoantigen for this autoimmune disease. BIP(78K) is also able to  
CC selectively proliferate synovial T-cells from patients with rheumatoid  
CC arthritis. BIP(78K) or peptides derived from the protein are useful as  
CC reagents to indicate the presence of rheumatoid arthritis, and can be  
CC used in prognostic or diagnostic tests of body fluids for rheumatoid  
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western  
CC blotting. The protein or the cDNA encoding it can also be used to test  
CC for rheumatoid arthritis by detecting antibodies to the protein.  
CC BIP(78K), its peptides and polynucleotides are also useful  
CC therapeutically.

XX Sequence 1917 BP; 577 A; 419 C; 487 G; 434 T; 0 other;

Query Match 100.0%; Score 1917; DB 21; Length 1917;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1917; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaagaggaacaaagagagagcgtgagcaggtggtcggcatcgacctggagcacc 60  
DB 1 atggaagaggaacaaagagagcgtgagcaggtggtcggcatcgacctggagcacc 60  
QY 61 taccctcgctcggtgttcaagaagcgccgctgtagatcatcgccaagatcaaggc 120  
DB 61 taccctcgctcggtgttcaagaagcgccgctgtagatcatcgccaagatcaaggc 120  
QY 121 aaccgcatcagccgctctatgtcgcttcactcctgaaggagagctctgattggcgt 180  
DB 121 aaccgcatcagccgctctatgtcgcttcactcctgaaggagagctctgattggcgt 180  
QY 181 gccgcaagaacacgctcaacctccaaccccgagaaacggtctcttgcgcgaagcgctc 240  
DB 181 gccgcaagaacacgctcaacctccaaccccgagaaacggtctcttgcgcgaagcgctc 240  
QY 241 atcgccgcaagctggaatgacccgctgtgtagagagagatcaagaatctcttgcgttcaag 300  
DB 241 atcgccgcaagctggaatgacccgctgtgtagagagagatcaagaatctcttgcgttcaag 300  
QY 301 gtggtctgaagaagaataaacaatacaatcaatgtatattgtaggtgtgagcaacaag 360  
DB 301 gtggtctgaagaagaataaacaatacaatcaatgtatattgtaggtgtgagcaacaag 360  
QY 361 acaattgctcctcgaagaataatctgcacatggttctcaataaagaaagaaacgcgtgag 420  
DB 361 acaattgctcctcgaagaataatctgcacatggttctcaataaagaaagaaacgcgtgag 420  
QY 421 gctaatctggagaagaaggttaaccaatgtagtcttactgtacagacctatttaagat 480  
DB 421 gctaatctggagaagaaggttaaccaatgtagtcttactgtacagacctatttaagat 480

QY 481 gcccaagccaaagcaaccaaagaacgctgagacatttgctgctaaatgtatagagatc 540  
DB 481 gcccaagccaaagcaaccaaagaacgctgagacatttgctgctaaatgtatagagatc 540  
QY 541 atcaacagagctacagcagctgctatgtctctatgacctgataagaagagggagagag 600  
DB 541 atcaacagagctacagcagctgctatgtctctatgacctgataagaagagggagagag 600  
QY 601 aacatcctgtgttgaacctgggtgaggaaccttcgagtgtctctccaccattgac 660  
DB 601 aacatcctgtgttgaacctgggtgaggaaccttcgagtgtctctccaccattgac 660  
QY 661 aagtgctctcgaagcttgagccacaacgaagagatcatctcgtgtgagagacctt 720  
DB 661 aagtgctctcgaagcttgagccacaacgaagagatcatctcgtgtgagagacctt 720  
QY 721 gaccagcgtgtcattgaaacctcaatcaactgtatacaaaagaagacggtcaaatgtc 780  
DB 721 gaccagcgtgtcattgaaacctcaatcaactgtatacaaaagaagacggtcaaatgtc 780  
QY 781 aggaagacataatagctgtgagaaatctcgccgagagtagaagaagccaaagggcc 840  
DB 781 aggaagacataatagctgtgagaaatctcgccgagagtagaagaagccaaagggcc 840  
QY 841 ctgctctcagcatcaagaacgaagatgaaatgagctccctcatgagagagagacctt 900  
DB 841 ctgctctcagcatcaagaacgaagatgaaatgagctccctcatgagagagagacctt 900  
QY 901 tctgagacacctgactcgagccaaatctgaaagatcaacatgatactgttcggtctact 960  
DB 901 tctgagacacctgactcgagccaaatctgaaagatcaacatgatactgttcggtctact 960  
QY 961 atgaagccgctccaagaagaggttgaaagatctcgatttgaaagagctgatatgagaa 1020  
DB 961 atgaagccgctccaagaagaggttgaaagatctcgatttgaaagagctgatatgagaa 1020  
QY 1021 atgtctctgtggtgctgactcgaaatctcaaaagatcaagaacctggttaaaagttc 1080  
DB 1021 atgtctctgtggtgctgactcgaaatctcaaaagatcaagaacctggttaaaagttc 1080  
QY 1081 tcaatgcaagaagacaatccctgtgcatcaaaagatgaagctgtagcgtatgtgtc 1140  
DB 1081 tcaatgcaagaagacaatccctgtgcatcaaaagatgaagctgtagcgtatgtgtc 1140  
QY 1141 gctgtcaagcctggtgtgtctctgtgtgatacaagaatcaagctgacgtgactgtgat 1200  
DB 1141 gctgtcaagcctggtgtgtctctgtgtgatacaagaatcaagctgacgtgactgtgat 1200  
QY 1201 gtaatgccccttaacacttgatataagaaactgtggaaggttcaatgacaaactgattcca 1260  
DB 1201 gtaatgccccttaacacttgatataagaaactgtggaaggttcaatgacaaactgattcca 1260  
QY 1261 aggaacaagtggtgtctcaacaagaagttcacaatcttctctacagctcttgataatcaa 1320  
DB 1261 aggaacaagtggtgtctcaacaagaagttcacaatcttctctacagctcttgataatcaa 1320  
QY 1321 ccaactgttaaatcaagagctatgaaaggtgaaagaccccttgacaaaagaacatcatctt 1380  
DB 1321 ccaactgttaaatcaagagctatgaaaggtgaaagaccccttgacaaaagaacatcatctt 1380  
QY 1381 ctggtgtaattgtaactgtaactgaaatccctcgtctcctgtggtggtcccaacagattgaa 1440  
DB 1381 ctggtgtaattgtaactgtaactgaaatccctcgtctcctgtggtggtcccaacagattgaa 1440  
QY 1441 gtcaacttgagatagatgtgaatggtatctctcgaatggaagctgaaagaaaggttaca 1500  
DB 1441 gtcaacttgagatagatgtgaatggtatctctcgaatggaagctgaaagaaaggttaca 1500  
QY 1501 gggaacaaaataaagatcaaatcaatcaatgaaacgaagatggcgtgaaacctgaaagaaatc 1560  
DB 1501 gggaacaaaataaagatcaaatcaatcaatgaaacgaagatggcgtgaaacctgaaagaaatc 1560



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 ; Search time 11004.8 Seconds  
(without alignments)  
2873.749 Million cell updates/sec

Title: US-09-806-955-3

Perfect score: 1917  
Sequence: 1 atgagagagagacaagaaga.....tcacccacacacacacacac 1917

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

1 number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenBank:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_cm:\*
- 5: gb\_ov:\*
- 6: gb\_pal:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_om:\*
- 20: em\_or:\*
- 21: em\_ov:\*
- 22: em\_pat:\*
- 23: em\_ph:\*
- 24: em\_pl:\*
- 25: em\_ro:\*
- 26: em\_sts:\*
- 27: em\_sy:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htgo\_hum:\*
- 31: em\_htgo\_inv:\*
- 32: em\_htgo\_rod:\*
- 33: em\_htg\_hum:\*
- 34: em\_htg\_inv:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1917	100.0	1917	9	AF188611	AF188611 Homo sapi
2	1894	98.8	1965	9	AF216292	AF216292 Homo sapi
3	1894	98.8	2007	9	HS4271729	AJ2717739 Homo sapi
4	1894	98.8	2639	6	AX014912	AX014912 Sequence
5	1866.8	97.4	2554	9	HSRABABP	X87949 H.sapiens m
6	1697.8	88.6	2453	4	SSGRP78	X92446 S.scrofa mr
7	1670	87.1	2282	10	MMBIPCHAP	AJ002387 Mus muscu
8	1665.8	86.9	181429	2	AL354702	AL354702 Homo sapi
9	1662	86.7	2334	10	CRUGRP	M17169 Chinese ham
10	1662	86.7	2408	10	MUSGRP784	D78645 Mouse mRNA
11	1636.4	85.4	159816	2	AC011026	AC011026 Homo sapi
12	1631.6	85.1	2383	10	RATBIP	M14050 Rat Immunog
13	1457	76.0	1667	10	BC005785	BC005785 Mus muscu
14	1426.4	74.4	2389	5	CHKGR78	M27260 Chicken 78-
15	1261	65.8	2470	5	XL062807	U62807 Xenopus lae
16	1232.2	64.3	2379	5	XL055069	U55069 Xenopus lae
17	1129.2	58.9	2643	3	AB016836	AB016836 Bombyx mo
18	1119.6	58.4	138532	2	AL139222	AL139222 Homo sapi
19	1077.6	56.2	2483	3	ACBIP	215041 A.californi
20	1068.8	55.8	3066	3	DROHSC3A	L01498 Drosophila
21	1031	53.8	7298	2	AC014024	AC014024 Drosophila
22	1003.6	52.4	2117	3	SDGR78	Y09500 S.domuncula
23	988.2	51.5	2152	3	AY046874	AY046874 Trichinel
24	938	48.9	1956	3	ECCGRPA	M63605 Echinosuccu
25	931.4	48.6	2156	3	ECCGRPA	M63604 Echinosuccu
26	925.8	48.3	2118	3	AF044412	AF044412 Schistos
27	851	44.4	2626	8	PMSLBIPRT	Z49764 Pseudotsuga
28	843.6	44.0	2345	8	NTBLP4	X60057 Nicotiana t
29	836.8	43.7	3859	8	PCGRP78G	X75673 P.cinamom
30	835.2	43.6	2007	8	CAV295617	AJ295617 Corylus a
31	834.8	43.5	2273	8	NTBLP5	X60058 Nicotiana t
32	834.4	43.5	2180	8	AF031241	AF031241 glycine m
33	834.4	43.5	2365	8	ZM058209	U58209 Zea mays 1u
34	827.2	43.2	2332	8	CSA249329	AJ249329 Cucumis s
35	827.2	43.1	2040	8	SDU12020	AJ312020 Scherffel
36	826.2	43.1	2410	8	SPIHSC70A	L23551 Spinacia ol
37	824.8	43.0	2232	8	TOMBIPGRBC	L08830 Tomato BIP
38	823.2	42.9	2338	8	ZM058208	U58208 Zea mays 1u
39	821.6	42.9	2417	8	AF006825	AF006825 Oryza sat
40	816.8	42.6	2569	8	YL063136	U63136 Yarrowia 11
41	810	42.3	32814	8	SPAC22A12	Z99295 S.pombe chr
42	805.6	42.0	2738	8	AF110397	AF110397 Toxoplas
43	803.8	41.9	2028	8	BLVHSPHAA	L32165 Hordeum vul
44	803.4	41.9	2220	8	ATHBIPA	D84414 Arabidopsis
45	799	41.7	2194	8	PTU29675	U29675 Phaeodactyl

## ALIGNMENTS

RESULT 1  
AF188611  
LOCUS AF188611 1917 bp mRNA PRI 29-NOV-1999  
DEFINITION Homo sapiens BIP protein (HSPB5) mRNA, partial cds.  
ACCESSION AF188611  
VERSION AF188611.1 GI:6470149

## KEYWORDS

human.  
Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 1917)

## AUTHORS

File,M.S.

## TITLE

Unpublished

## JOURNAL

2 (bases 1 to 1917)

## REFERENCE

File,M.S.

## AUTHORS

Direct Submission

## TITLE

Submitted (21-SEP-1999) Molecular Immunogenetics, King's College,  
Guy's Hospital, St. Thomas' Street, London SE1 9RT, UK

## SUMMARIES



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:27:51 ; Search time 18.76 Seconds

(without alignments)  
1237.145 Million cell updates/sec

Title: US-09-806-955-2

Perfect score: 3225

Sequence: 1 MEEDKEDVGVVIGIDLT.....SKYGSAGPPTEEDTAEL 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

al number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3216	99.7	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	99.5	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	99.4	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	99.2	654	1 GR78_RAT	P06761 rattus norv
5	3168	98.2	652	1 GR78_CHICK	G05053 gallus gall
6	3069	95.2	658	1 GR78_XENLA	G01883 xenopus lae
7	2722.5	84.4	667	1 GR78_APLCA	Q16956 aplysia cal
8	2708	84.0	656	1 GR78_DROME	P29844 drosophila
9	2594.5	73.0	661	1 HS7C_GABEL	P27420 caenorhabdi
10	2354.5	72.8	668	1 BIP5_TOBAC	Q03684 nicotiana t
11	2346.5	72.8	668	1 BIP5_TOBAC	Q03685 nicotiana t
12	2332.5	72.3	663	1 BIP3_MAIZE	Q24581 zea mays (m
13	2332.5	72.2	668	1 BIP_LYCES	P49118 lycopersico
14	2324.5	72.1	668	1 BIP_SPIOL	Q42424 spinacia ol
15	2322.5	72.0	663	1 BIP2_MAIZE	P24067 zea mays (m
16	2298.5	71.3	669	1 BIP1_ARATH	G01KX3 arabidopsis
17	2296	71.2	668	1 BIP2_ARATH	G039043 arabidopsis
18	2227.5	69.1	662	1 GR78_NEUCR	P78655 neurospora
19	2216	68.7	663	1 GR78_SCHPO	P26604 schizosacch
20	2191	67.9	670	1 GR78_YARLI	G09110 yarrowia li
21	2173	67.4	682	1 GR78_YEAST	P16474 saccharomyc
22	2100	65.1	655	1 GR78_PLAFO	Q05866 plasmodium
23	2085.5	64.7	641	1 HS71_YEAST	P10591 saccharomyc
24	2083	64.6	646	1 HS7C_HUMAN	P11142 homo sapien
25	2083	64.6	646	1 HS7C_MOUSE	P08108 mus musculu
26	2079	64.5	650	1 HS7C_BOVIN	P19120 bos taurus
27	2078	64.4	644	1 HS71_PICAN	P23421 pichia angu
28	2076.5	64.4	640	1 HS7C_DICDI	P26415 dictyosteli
29	2074.5	64.3	655	1 HS71_CANAL	P13797 candida alb
30	2074	64.3	646	1 HS7C_CRIGR	P19378 cricetus
31	2072	64.2	638	1 HS72_YEAST	P10592 saccharomyc
32	2065.5	64.0	679	1 GR78_KLULA	P22010 kluyveromyc
33	2065	64.0	648	1 HS71_PUCGR	G01877 puccinia gr

34	2064	64.0	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	64.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.9	646	1 HS70_NEUCR	Q01233 neurospora
37	2060	63.9	643	1 HS71_SCHPO	Q10265 schizosacch
38	2059	63.8	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	63.7	651	1 HS7C_PETHY	P09189 petunia hyb
40	2052	63.6	652	1 HS7D_MANSE	G09639 manduca sex
41	2049	63.5	651	1 HS70_ONCMY	P08108 oncorhynch
42	2046	63.4	641	1 HS72_BOVIN	Q27965 bos taurus
43	2045	63.4	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	63.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	63.2	641	1 HS71_HUMAN	P08107 homo sapien

## ALIGNMENTS

RESULT	ID	GR78_HUMAN	STANDARD;	PRT;	654 AA.
AC	P11021	GR78_HUMAN	09NPFL		
DT	01-JUL-1989	(Rel. 11, Created)			
DT	20-AUG-2001	(Rel. 40, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN				
DE	HEAVY CHAIN BINDING PROTEIN) (BIP) (ENDOPLASMIC RETICULUM LUMENAL CA2+				
DE	BINDING PROTEIN GRP78).				
GN	HSP45 OR GRP78.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-86283347; Pubmed-2840249;				
RA	Ting J., Lee A.S.;				
RT	*Human gene encoding the 78,000-dalton glucose-regulated protein and				
RT	its pseudogene: structure, conservation, and regulation.*;				
RL	DNA 7:275-286(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Fibroblast;				
RA	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;				
RT	*Grp78 is involved in the quality control of the LDL-receptor.*;				
RL	submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Bernandez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;				
RT	*Sequence differences between human grp78/BIP isolated from HeLa cells				
RT	and previously reported human sequences.*;				
RL	submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 1-25 FROM N.A.				
RX	MEDLINE-93126087; Pubmed-1480470;				
RA	Chao C.C.K., Lin-Chao S.;				
RT	*A direct-repeat sequence of the human BIP gene is required for				
RT	A23187-mediated inducibility and an inducible nuclear factor				
RT	binding.*;				
RL	Nucleic Acids Res. 20:6481-6485(1992).				
RN	[6]				
RP	SEQUENCE OF 22-38.				
RC	TISSUE-Breast carcinoma;				
RX	MEDLINE-97295304; Pubmed-9150946;				
RA	Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.;				
RT	*Two-dimensional electrophoretic analysis of human breast carcinoma				
RT	proteins: mapping of proteins that bind to the SH3 domain of mixed				
RT	lineage kinase MLK2.*;				

```

RL Electrophoresis 18:588-598(1997).
RN [7]
RP SEQUENCE OF 19-40.
RC TISSUE-Colon carcinoma; PubMed=9150948;
RX MEDLINE=97295306; PubMed=9150948;
RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins."
RL Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M19645; AAA52614.1; -
DR EMBL: X87949; CAAG1201.1; -
DR EMBL: AJ271729; CAB7135.1; -
DR EMBL: AF216292; AAF42836.1; -
DR EMBL: X59969; CAA42595.1; -
DR PIR: A29821; A29821.
DR HSP: P19120; IATR.
DR SWISS-2DPAGE: P11021; HUMAN.
DR MIM: 138120; -
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET.1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR ATP-binding; Endoplasmic reticulum; signal.
FT SIGNAL 1 18
FT CHAIN 19 654
FT SITE 651 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
FT CONFLICT 457 457 R -> N (IN REF. 1 AND 2).
SQ SEQUENCE 654 AA; 72333 MW; 59B7DB85BC32A00 CRC64;

```

```

Query Match 99.7%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 2.5e-152;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 EEDKKDDVTVVIGIDGTYSCVGFKNRVELIANDQGNRTTPSVATTPREERLIGA 61
DB 20 EEDKKDDVTVVIGIDGTYSCVGFKNRVELIANDQGNRTTPSVATTPREERLIGA 79
QY 62 AKNQLTSNPNVYFADKRLIGRTWNPVQODIKFLPFYVEKRTKPYIOVDIGGQRT 121
DB 80 AKNQLTSNPNVYFADKRLIGRTWNPVQODIKFLPFYVEKRTKPYIOVDIGGQRT 139
QY 122 FAPBEISAVVLTQKMETAEAYLGKVTAVVTVPAYFNDQROATKDACTIAGLVNRII 181
DB 140 FAPBEISAVVLTQKMETAEAYLGKVTAVVTVPAYFNDQROATKDACTIAGLVNRII 199
QY 182 NEPTAAIAVYGDREGEKNILVFDLGGSTFDVSLTTINGVEVVAATGDLHGGEDD 241
DB 200 NEPTAAIAVYGDREGEKNILVFDLGGSTFDVSLTTINGVEVVAATGDLHGGEDD 259
QY 242 ORVMEHFILYKKKTKGDKVRKONRAVOKLRREYERAKRALLSSOHARIEIESFEDEDS 301
DB 260 ORVMEHFILYKKKTKGDKVRKONRAVOKLRREYERAKRALLSSOHARIEIESFEDEDS 319

```

```

QY 302 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVCSTRIPIQOLKEEF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVCSTRIPIQOLKEEF 379
QY 362 NKEPSRGINPDEAVAYGAOVAGVLSGDDTGDVLDVCPVLUGIEFVGVMKILPR 421
DB 380 NKEPSRGINPDEAVAYGAOVAGVLSGDDTGDVLDVCPVLUGIEFVGVMKILPR 439
QY 422 NVVPPKKSQIFSTSDNDPVTIKYEGEERPLTXONHILGTFDILGIPAPRGVQIIV 481
DB 440 NVVPPKKSQIFSTSDNDPVTIKYEGEERPLTXONHILGTFDILGIPAPRGVQIIV 499
QY 482 TEIDVNGILRYAEDKGGGNKKITITNDONRLPPEIERVNDQAEFEEDKKLKERT 541
DB 500 TEIDVNGILRYAEDKGGGNKKITITNDONRLPPEIERVNDQAEFEEDKKLKERT 559
QY 542 DTRNELESYAYSILKNOIGKEKLGKLSSEDEKTEMKAVEKIEWLESQDDADIEDPKK 601
DB 560 DTRNELESYAYSILKNOIGKEKLGKLSSEDEKTEMKAVEKIEWLESQDDADIEDPKK 619
QY 602 KKELEIVOPITISKLYGSAGPPTGEEDTAE 632
DB 620 KKELEIVOPITISKLYGSAGPPTGEEDTAE 650

```

```

RESULT 2
GR78_MESAU STANDARD; PRT: 654 AA.
ID GR78_MESAU
AC P07823;
DT 01-AUG-1988 (rel. 08; Created)
DT 01-AUG-1988 (rel. 08; Last sequence update)
DT 01-NOV-1997 (rel. 35; Last annotation update)
DE 78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN
DE HEAVY CHAIN BINDING PROTEIN) (BIP).
GN HSP45 OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87305586; PubMed=3623104;
RA Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat."
RL Gene 53:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17169; AAA51448.1; -
DR PIR: A27414; A27414.
DR HSP: P19120; IATR.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET.1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR ATP-binding; Endoplasmic reticulum; signal.
KW

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:23:14 ; Search time 23.81 Seconds  
(without alignments)  
598.261 Million cell updates/sec

Title: US-09-806-955-2

Perfect score: 3225  
Sequence: 1 MEEDKKEDVGVIGIDLGT.....SKLYSAGPPPTGEEDTAEI 633

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 21252 seqs, 22503292 residues

al number of hits satisfying chosen parameters: 21252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	99.5	654	1 US-08-441-139-11	Sequence 11, Appl
2	3168	98.2	666	1 US-08-441-139-16	Sequence 16, Appl
3	2173	67.4	682	1 US-08-441-139-2	Sequence 2, Appl
4	2156	66.9	663	1 US-08-441-139-7	Sequence 7, Appl
5	2077	64.4	646	1 US-08-441-139-14	Sequence 14, Appl
6	2065.5	64.0	679	1 US-08-441-139-5	Sequence 5, Appl
7	1938	60.1	643	4 US-08-797-358B-3	Sequence 3, Appl
8	1547.5	48.0	679	1 US-08-214-583-2	Sequence 2, Appl
9	1474	45.7	641	1 US-08-441-139-4	Sequence 4, Appl
10	1461.5	45.3	649	4 US-09-066-047-5	Sequence 5, Appl
11	1444	44.8	607	2 US-08-472-534-5	Sequence 5, Appl
12	1303.5	40.4	600	6 5240706-1	Patent No. 5240706
13	1140.5	35.4	339	2 US-08-928-692-52	Sequence 52, Appl
14	941	29.2	187	6 5196523-13	Patent No. 5196523
15	824	25.6	168	1 US-08-441-139-10	Sequence 10, Appl
16	818.5	25.4	471	1 US-08-257-073-7	Sequence 7, Appl
17	750.5	23.3	471	1 US-08-203-905B-2	Sequence 2, Appl
18	726.5	22.5	472	1 US-08-203-905B-14	Sequence 14, Appl
19	642.5	19.9	129	6 5196523-10	Patent No. 5196523
20	607.5	18.8	999	2 US-08-770-301A-3	Sequence 3, Appl
21	607.5	18.8	999	2 US-09-175-581-3	Sequence 3, Appl
22	598	18.5	999	2 US-08-770-301A-1	Sequence 1, Appl
23	598	18.5	999	2 US-09-175-581-1	Sequence 1, Appl
24	528.5	16.4	560	2 US-08-928-692-53	Sequence 53, Appl
25	437.5	13.6	374	2 US-08-928-692-51	Sequence 51, Appl
26	402	12.5	77	6 5196523-7	Patent No. 5196523
27	395	12.2	79	6 5196523-11	Patent No. 5196523

28	367.5	11.4	599	4 US-09-080-983-9	Sequence 9, Appl
29	358	11.1	549	2 US-08-770-544-6	Sequence 6, Appl
30	345.5	10.7	80	1 US-08-464-164-4	Sequence 4, Appl
31	345.5	10.7	80	1 US-08-338-057-4	Sequence 4, Appl
32	345.5	10.7	80	2 US-08-668-416-4	Sequence 4, Appl
33	323	10.0	554	5 PCT-US94-06430-7	Sequence 7, Appl
34	320	9.9	554	4 US-08-591-468-7	Sequence 7, Appl
35	267.5	8.3	714	2 US-08-472-534-3	Sequence 7, Appl
36	264	8.2	56	6 5196523-12	Patent No. 5196523
37	230	7.1	37	6 5196523-8	Patent No. 5196523
38	211	6.5	42	6 5196523-15	Patent No. 5196523
39	192	6.0	38	6 5196523-9	Patent No. 5196523
40	168.5	5.2	1786	4 US-08-973-462-8	Sequence 8, Appl
41	147	4.6	693	4 US-08-235-836C-68	Sequence 68, Appl
42	143.5	4.4	1805	1 US-07-853-913-2	Sequence 2, Appl
43	142.5	4.4	2285	4 US-09-308-375-2	Sequence 2, Appl
44	138	4.3	700	1 US-07-720-589-2	Sequence 2, Appl
45	138	4.3	700	2 US-08-785-190-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-441-139-11  
Sequence 11, Application US/08441139  
Patent No. 5773245  
GENERAL INFORMATION:  
APPLICANT: Wittrup, Dr. Karl D.  
APPLICANT: Robinson, Anne S.  
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF  
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,139  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/089,997  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8646  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 654 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-441-139-11  
Query Match 99.5%, Score 3209, DB 11, Length 654:  
Best Local Similarity 99.7%, Pred. No. 4,1e-259:  
Matches 629, Conservative 2, Mismatches 0, Indels 0, Gaps 0:

```

QY      2 EEDKKEDVGVGIDLTGTTSCVGFKNRVEIILANDQNRITPSYAFTEGGERLIGDA 61
      20 EEDKKEDVGVGIDLTGTTSCVGFKNRVEIILANDQNRITPSYAFTEGGERLIGDA 79
      62 AKNOLTSNPENTVPDAKRLIGRTWNDSVOODIKFLPFKVEKKTKPYIOVDIGCGGCTKT 121
      80 AKNOLTSNPENTVPDAKRLIGRTWNDSVOODIKFLPFKVEKKTKPYIOVDIGCGGCTKT 139
      122 FAPBEISAMVLTMMKEAEYLGKRYHVVTPAYPANDQROATKAGTIAGLNVRIT 181
      140 FAPBEISAMVLTMMKEAEYLGKRYHVVTPAYPANDQROATKAGTIAGLNVRIT 199
      182 NEPTAAIAVGLDKREGEKNILVFDLGGTFEDVSLTIDNGVEVATNGDTHLGGEDFD 241
      200 NEPTAAIAVGLDKREGEKNILVFDLGGTFEDVSLTIDNGVEVATNGDTHLGGEDFD 259
      242 QRVMEHFILYKKKKTGKDVKKRRAVQKLRREVEKAKRALSQHOARIEIESFEEDGDFS 301
      260 QRVMEHFILYKKKKTGKDVKKRRAVQKLRREVEKAKRALSQHOARIEIESFEEDGDFS 319
      302 ETLTFRKFEELNDLFRSTMKPVQKYLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 361
      320 ETLTFRKFEELNDLFRSTMKPVQKYLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 379
      362 NGKEPSKGINPDEAVAVGAAGVAVLSGDDTGDVLDVCPPLTIGETVGVMTKLIPR 421
      380 NGKEPSKGINPDEAVAVGAAGVAVLSGDDTGDVLDVCPPLTIGETVGVMTKLIPR 439
      422 NTVPPTKKSQIFSTASDNQPTVIKYEGERPLTKDNHLLGTDLGIPAPRGVQIEV 481
      440 NTVPPTKKSQIFSTASDNQPTVIKYEGERPLTKDNHLLGTDLGIPAPRGVQIEV 499
      482 TFEIDVNGILRVTAEDKGTGNKNTTTNDQNLTPREIRMVNDKFAEEDKKLKERI 541
      500 TFEIDVNGILRVTAEDKGTGNKNTTTNDQNLTPREIRMVNDKFAEEDKKLKERI 559
      542 DTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEKIEMLESHODADIEDPKAK 601
      560 DTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEKIEMLESHODADIEDPKAK 619
      602 KKELEEVOPITISKLYGSAGPPTGGEEDTAE 632
      620 KKELEEVOPITISKLYGSAGPPTGGEEDTAE 650
      DB
      RESULT 2
      US-08-441-139-16
      ; Sequence 16, Application US/08441139
      ; Patent No. 5773245
      ; GENERAL INFORMATION:
      ; APPLICANT: Mittrud, Dr. Karl D.
      ; APPLICANT: Robinson, Anne S.
      ; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
      ; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
      ; NUMBER OF SEQUENCES: 20
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
      ; STREET: 400 Garden City Plaza
      ; City: Garden City
      ; STATE: NY
      ; COUNTRY: USA
      ; ZIP: 11530
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; OPERATING SYSTEM: IBM PC compatible
      ; SOFTWARE: Patent Release #1.0, Version #1.25
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/441,139
      ; FILING DATE: 15-MAY-1995
      ; CLASSIFICATION: 435
      ; PRIORITY APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digillo, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-139-16

```

```

Query Match          98.2%  Score 3168:  DB 1:  Length 666:
Best Local Similarity 97.8%  Pred. No. 1.1e-255:
Matches 617:  Conservative 11:  Mismatches 3:  Indels 0:  Gaps 0:

```

```

QY      2 EEDKKEDVGVGIDLTGTTSCVGFKNRVEIILANDQNRITPSYAFTEGGERLIGDA 61
      32 DEKKEDVGVGIDLTGTTSCVGFKNRVEIILANDQNRITPSYAFTEGGERLIGDA 91
      62 AKNOLTSNPENTVPDAKRLIGRTWNDSVOODIKFLPFKVEKKTKPYIOVDIGCGGCTKT 121
      92 AKNOLTSNPENTVPDAKRLIGRTWNDSVOODIKFLPFKVEKKAKPHIOVDVGGGCTKT 151
      122 FAPBEISAMVLTMMKEAEYLGKRYHVVTPAYPANDQROATKAGTIAGLNVRIT 181
      152 FAPBEISAMVLTMMKEAEYLGKRYHVVTPAYPANDQROATKAGTIAGLNVRIT 211
      182 NEPTAAIAVGLDKREGEKNILVFDLGGTFEDVSLTIDNGVEVATNGDTHLGGEDFD 241
      212 NEPTAAIAVGLDKREGEKNILVFDLGGTFEDVSLTIDNGVEVATNGDTHLGGEDFD 271
      242 QRVMEHFILYKKKKTGKDVKKRRAVQKLRREVEKAKRALSQHOARIEIESFEEDGDFS 301
      272 QRVMEHFILYKKKKTGKDVKKRRAVQKLRREVEKAKRALSQHOARIEIESFEEDGDFS 331
      302 ETLTFRKFEELNDLFRSTMKPVQKYLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 361
      332 ETLTFRKFEELNDLFRSTMKPVQKYLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 391
      362 NGKEPSKGINPDEAVAVGAAGVAVLSGDDTGDVLDVCPPLTIGETVGVMTKLIPR 421
      392 NGKEPSKGINPDEAVAVGAAGVAVLSGDDTGDVLDVCPPLTIGETVGVMTKLIPR 451
      422 NTVPPTKKSQIFSTASDNQPTVIKYEGERPLTKDNHLLGTDLGIPAPRGVQIEV 481
      452 NTVPPTKKSQIFSTASDNQPTVIKYEGERPLTKDNHLLGTDLGIPAPRGVQIEV 511
      482 TFEIDVNGILRVTAEDKGTGNKNTTTNDQNLTPREIRMVNDKFAEEDKKLKERI 541
      512 TFEIDVNGILRVTAEDKGTGNKNTTTNDQNLTPREIRMVNDKFAEEDKKLKERI 571
      542 DTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEKIEMLESHODADIEDPKAK 601
      572 DTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEKIEMLESHODADIEDPKAK 631
      602 KKELEEVOPITISKLYGSAGPPTGGEEDTAE 632
      632 KKELEEVOPITISKLYGSAGPPTGGEEDTAE 662
      DB
      RESULT 3
      US-08-441-139-2
      ; Sequence 2, Application US/08441139
      ; Patent No. 5773245

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:22:43 ; Search time 43 Seconds  
(without alignments)  
1090.428 Million cell updates/sec

Title: US-09-806-955-2

Perfect score: 3225  
Sequence: 1 MEEDKKEDVGTVCIDIGT.....SKLYGSGAPPTGEEDTAEI 633

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

al number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

\_A\_Geneseq\_1101:\*

1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3225	100.0	633	21	AA90694
2	3225	100.0	639	21	AA90693
3	3209	99.5	654	15	AA953075
4	3182.5	98.7	653	22	AA82533
5	3171	98.3	656	14	AA93934
6	3168	98.2	666	15	AA953076
7	2296	71.2	668	21	AA941381
8	2296	71.2	691	21	AA941382
9	2258.5	70.0	642	21	AA939285
10	2258.5	70.0	658	21	AA939284
11	2237	69.4	655	21	AA93912

12	2173	67.4	682	22	AA970749
13	2122.5	65.8	687	22	AA97075
14	2083	64.6	686	19	AA954364
15	2083	64.6	646	20	AA917407
16	2083	64.6	646	20	AA917408
17	2083	64.6	646	21	AA923649
18	2083	64.6	646	21	AA923651
19	2083	64.6	646	22	AA923653
20	2083	64.6	646	22	AA923658
21	2074.5	64.3	656	18	AA901638
22	2064	64.3	646	11	AA903927
23	2041	63.3	647	11	AA903928
24	2039.5	63.2	624	21	AA923252
25	2039.5	63.2	641	19	AA954349
26	2039.5	63.2	641	21	AA923652
27	2039.5	63.2	641	22	AA923654
28	2036.5	63.1	646	21	AA923654
29	2033.5	63.1	633	14	AA943002
30	2032.5	63.0	647	20	AA931380
31	2032.5	63.0	647	22	AA949099
32	2032	63.0	634	11	AA903930
33	2021	62.7	640	21	AA923653
34	2017.5	62.6	640	11	AA903929
35	2014	62.4	642	21	AA923650
36	2013	62.4	640	18	AA910065
37	2013	62.4	640	21	AA988408
38	2013	62.4	677	11	AA909418
39	2008	62.3	641	14	AA943004
40	2004.5	62.2	669	11	AA903925
41	1994.5	61.8	676	9	AA980088
42	1974	61.2	665	21	AA958386
43	1970	61.1	646	11	AA903931
44	1968.5	61.0	657	11	AA903926
45	1963.5	60.9	623	21	AA953605

#### ALIGNMENTS

RESULT 1

AA90694

AA90694 standard; Protein: 633 AA.

AC AA90694;

29-AUG-2000 (first entry)

DE 633 residue human immunoglobulin heavy binding protein, B1P(78KD).

XX Immunoglobulin heavy chain binding protein; B1P(78KD); chondrocyte;

KW autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.

XX Homo sapiens.

OS WO200021995-A1.

PN 20-APR-2000.

PD 08-OCT-1999; 99WO-GB03316.

PE 09-OCT-1998; 98GB-0022115.

PR (UNLO ) KINGS COLLEGE LONDON.

PA Panayl GS, Corrigall VM, Bodman-Smith MD, Fife MS, Lanchbury JS;

PI WPI: 2000-317942/27.

DR New human immunoglobulin heavy chain binding protein and encoding

XX polynucleotide, useful for diagnosis and treatment of rheumatoid

PT arthritis

PS Claim 3; Page 44-46; 53pp; English.

XX This sequence represents a human immunoglobulin heavy chain binding  
CC protein, BIP(78KD), having a 533 amino acid sequence. The invention  
CC also encompasses a BIP(78KD) protein of 639 amino acids (Y906933).  
CC The cDNA encoding BIP(78KD), also referred to as p78 in the  
CC specification, was isolated from human chondrocytes (the specialised  
CC cells of articular cartilage) and human chondrosarcoma cell lines. The  
CC BIP(78KD) cDNA of this invention contains a number of differences  
CC compared with the published sequence (Genbank accession number X87949)  
CC and has therefore has been deposited with Genbank with the accession  
CC number AF188611). These differences comprise 6 single nucleotide  
CC substitutions and a codon insertion, and result in three amino acid  
CC substitutions and an arginine insertion at position 834-836 of the  
CC protein. The BIP(78KD) proteins react with antibodies present in the  
CC serum of rheumatoid arthritis patients, and is therefore a putative  
CC autoantigen for this autoimmune disease. BIP(78KD) is also able to  
CC selectively proliferate synovial T-cells from patients with rheumatoid  
CC arthritis. BIP(78KD) or peptides derived from the protein are useful  
CC reagents to indicate the presence of rheumatoid arthritis, and can be  
CC used in prognostic or diagnostic tests of body fluids for rheumatoid  
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western  
CC blotting. The protein or the cDNA encoding it can also be used to test  
CC for rheumatoid arthritis by detecting antibodies to the protein.  
CC BIP(78KD), its peptides and polynucleotides are also useful  
CC therapeutically.

Sequence 633 AA;

Query Match	100.0%;	Score 3225;	DB 21;	Length 633;
Best Local Similarity	100.0%;	Pred. No. 2.3e-208;		
Matches 633; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy	1	MEODKEDVGVNVDIGDSTTSCVGVCKRKNNGVVEIANDQGNRTIPSVAAATPGEERLIGD	60
Dd	1	meedkkelvvgvlgldlgltvscvgtknngvveilandgnrtlpsvvaatlpegerlvgd	60
Qy	61	AAKNQITSNPENTYFDAKRLRGRTMNDPSPVOODIKFLPFKVEKTKPTUYOVODGGQOTK	120
Dd	61	aaKnqItSnPeNtYfDaKRLrGRtMNDpSPVoODiKfLpFkVeKtKpTUYoVoDgGqOTk	120
Qy	121	TFPAEESAMVLTMMKETAETAEATLCKKVTYHAAVTVPAYFNDARQATKAGTLAGLVNMTI	180
Dd	121	tFpaEeSaMvLtMmKtEaEtAeAtLcKkVtYhAAvTvPaYfNdARqATkAGtLaGtLaGlvNmTi	180
Qy	181	INEPAAALAVGLDKRBEKENILFEDJGGGFQDVSILTITNGVREYAVATGNPHLGSEDF	240
Dd	181	INepAAaLAvGldKRBeKeNiLfEdJgGgFqDvSiLtItNgVReYAvAtGnPhLgSeDf	240
Qy	241	DQRVMEHITKLYKKKCTGDKVAKDRAAVOKLREVEKAKRAISOHQARIESFSEGEDE	3000
Dd	241	dQrVMEhITkLyKkCtGdKvAKdRAaVoKLrEvEkAKrAiSoHqARiEsFSeGeDe	3000
Qy	301	SETLTRAKFEELNMDLEFRSTNKPVQKVLVEDSLDKSDIDEIVLGGSTRIPKIQULVKEF	3600
Dd	301	seLtTrAkFeElNmDlEfRstNkPvQkVlEvEdSlDkSdIdEiVlGgStRiPkiQulVkeF	3600
Qy	361	FNGKPSHGIMPDEAVANGAUVQAGVLSGODPTDVLVLDVCPRTLGIETVGGVMTRKLP	4200
Dd	361	fNgKpShGImPDeAvANgAuVqAGvLsGoDpTdVlVldVcPrtLgiEtVgGvMtrKlp	4200
Qy	421	RNTVPTPKKSQIFSTASPNQPTVITIKYEGSERPLTKCNHLLGTFDPLGIPAPGAVQIE	4800
Dd	421	rNtVpTpkKsQIFsTaSPnQpTvITiKyeGseRPltKcNhLlGtFdPlGiPaPgAvQie	4800
Qy	481	VTFEIDVNGILRVTAEDKGTKNNKKTITNDONRLTPBEETIERMVNDAEKFAEEDKILKER	5400
Dd	481	vTfEiDvNgIlRvTaEdKgtKnKktItNdOnrLtpBeEtIeRmVnDaEkFaEedKiLkeR	5400
Qy	541	IDTRNELESYAVSLKNOJGDKEXKLGKRLSSDKDITKMKAYBEKLEMLESHQODADIEFKA	6000
Dd	541	IdTrNeLeSyAvSlKnOjGdKExKlGkRlSSdKdITkMkAyBeKlEmLeShQoDaDiEfKa	6000

Qy	601	kkkeleeivopiisklygsacppptgeedtael	633
Db	601	kkkeleeivopisklygsacppptgeedtael	633

## RESULT 2

ID AAY90693 standard; Protein; 639 AA.

AC MAY90693;

DT 29-AUG-2000 (first entry)

DE 639 residue human immunoglobulin heavy binding protein, BiP(78KD).

KW Immunoglobulin heavy chain binding protein; BLP(78KD); chondrocyte; autointigen; rheumatoid arthritis; antiarthritis; antirheumatic; p78

OS Homo sapiens.

PN WO200021995-A1

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

PA (UNIO ) KINGS COLLEGE LONDON.

PI Panayl GS, Corrigan VM, Bodman-Smith MD, Flfe MS, Lanchbury JS;

DR WPI; 2000-317942/27

DR N-PSDB; AAA30792.

PT New human immunoglobulin heavy chain binding protein and encoding polynucleotide, useful for diagnosis and treatment of rheumatoid

PS Claim 3; Page 23; 53pp; English.

CC This sequence represents a human immunoglobulin heavy chain binding  
CC protein, B1P(78KD), having a 639 amino acid sequence. The invention  
CC also encompasses a B1P(78KD) protein of 633 amino acids (190654).  
CC The cDNA encoding B1P(78KD), also referred to as p78 in the  
CC specification, was isolated from human chondrocytes (the specialised  
CC cells of articular cartilage) and human chondrosarcoma cell lines. The  
CC B1P(78KD) cDNA of this invention contains a number of differences  
CC compared with the published sequence (Genbank accession number X87949)  
CC and has therefore been deposited with Genbank with the accession  
CC number A118611). These differences comprise 6 single nucleotide  
CC substitutions and a codon insertion, and result in three amino acid  
CC substitutions and an arginine insertion at position 834-836 of the  
CC protein. The B1P(78KD) proteins react with antibodies present in the  
CC serum of rheumatoid arthritis patients, and is therefore a putative  
CC autoantigen for this autoimmune disease. B1P(78KD) is also able to  
CC selectively proliferate synovial T-cells from patients with rheumatoid  
CC arthritis. B1P(78KD) or peptides derived from the protein are useful as  
CC reagents to indicate the presence of rheumatoid arthritis, and can be  
CC used in prognostic or diagnostic tests of body fluids for rheumatoid  
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western  
CC blotting. The protein or the cDNA encoding it can also be used to test  
CC for rheumatoid arthritis by detecting antibodies to the protein.  
CC B1P(78KD), its peptides and polynucleotides are also useful  
CC therapeutically.

**SQ** Sequence 639 AA;

Query Match	100.08;	Score 3225;	DB 21;	Length 639;
Best Local Similarity	100.08;	Pred. No. 2.3e-208;		
Matches 633;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:27:27 : Search time 50.77 Seconds  
(without alignments)  
1823.722 Million cell updates/sec

Title: US-09-806-955-2  
Perfect score: 3225  
Sequence: 1 MEEDKKEDVGVGIDLTGTT.....SKLYSGAGPPPTGEPTAEL 633

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
a1 number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database :  
1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP\_mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3225	100.0	639	4	Q9UK02
2	3216	99.7	654	4	Q9NPF1
3	3199	99.2	655	11	Q9DC41
4	2994	92.8	658	13	Q91688
5	2725	84.5	658	5	Q76180
6	2717	84.2	656	5	Q9VYU3
7	2607.5	80.9	656	5	Q62568
8	2555	79.2	657	5	Q02085
9	2534	78.6	651	5	Q24798
10	2525.5	78.3	649	5	Q24895
11	2460	76.3	648	5	Q45038
12	2344.5	72.7	665	10	Q9M4E8
13	2325	72.1	668	10	Q22639
14	2310.5	71.6	675	10	Q40924
15	2305.5	71.5	663	10	Q24182
16	2291	71.0	668	10	Q9FS17
17	2283	70.8	672	3	Q14453
18	2276	70.6	672	3	Q13280
19	2273	70.5	668	5	Q9U540

20	2263	70.2	669	3	Q9C1C1	Q9C1C1 cronartium
21	2261	70.1	642	5	Q61001	Q61001 toxoplasma
22	2244.5	69.6	672	3	Q9UWE3	Q9UWE3 aspergillus
23	2243	69.6	659	10	Q41074	Q41074 phaeodactyl
24	2237	69.4	655	10	Q04022	Q04022 arabidopsis
25	2234.5	69.3	664	10	Q39830	Q39830 glycine max
26	2231	69.2	658	10	Q12752	Q12752 phytophor
27	2226.5	69.0	701	5	Q24928	Q24928 elmeria ten
28	2214.5	68.7	608	10	Q40058	Q40058 hordium vul
29	2200	68.2	666	10	Q39804	Q39804 glycine max
30	2188.5	67.9	655	3	P87028	P87028 pneumocysti
31	2156.5	66.9	665	3	Q9HG01	Q9HG01 pichia angu
32	2139.5	66.3	652	3	Q00053	Q00053 pneumocysti
33	2116.5	65.6	652	5	Q07615	Q07615 plasmodium
34	2108	65.4	651	3	Q9UVM1	Q9UVM1 rhizopus st
35	2093	64.9	659	5	Q9XZJ2	Q9XZJ2 crassostrea
36	2080	64.5	642	3	Q9UVM0	Q9UVM0 rhizopus st
37	2079	64.5	647	3	Q59855	Q59855 schizosacch
38	2074.5	64.3	648	10	Q40980	Q40980 pisum sativ
39	2074	64.3	661	5	Q77164	Q77164 entamoeba h
40	2069	64.2	646	13	Q73885	Q73885 gallus gall
41	2063.5	64.0	647	3	Q94104	Q94104 pneumocysti
42	2062.5	64.0	645	3	Q94106	Q94106 pneumocysti
43	2057	63.8	652	5	Q26924	Q26924 trypanosoma
44	2055.5	63.7	647	10	Q9M4E6	Q9M4E6 cucumis sat
45	2055.5	63.7	653	5	Q26733	Q26733 trypanosoma

## ALIGNMENTS

RESULT 1  
ID Q9UK02 PRELIMINARY: PRT: 639 AA.

AC Q9UK02: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE BIP PROTEIN (FRAGMENT).  
GN HSPA5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-ARTICULAR CARTILAGE;  
RA File M.S.;  
RT "Bip: an Autoantigen associated with Rheumatoid Arthritis.";  
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF188611; AAF13605.1; -  
DR HSRP: P19120; 3HSC.  
DR InterPro: IPR001023; HSP70.  
DR Pfam: PF00012; HSP70.1.  
DR PRINTS: PR00301; HEATSHOCK70.  
DR PROSITE: PS00297; HSP70.1; 1.  
DR PROSITE: PS00329; HSP70.2; 1.  
DR PROSITE: PS01036; HSP70.3; 1.  
FT NON\_TER 639  
SQ SEQUENCE 639 AA: 70931 MW: 3421208FP128FD5D CRC64:

Query Match 100.0%; Score 3225; DB 4; Length 639;  
Best Local Similarity 100.0%; Pred. No. 3.8e-156;  
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEDKKEDVGVGIDLTGTTSCVGFKNGRVEIITANDOGNRTSPSYAFTEGSRLLGD 60  
DB 1 MEEDKKEDVGVGIDLTGTTSCVGFKNGRVEIITANDOGNRTSPSYAFTEGSRLLGD 60  
QY 61 AAKNOLTSNPENTVFDAKRLIGRTWNDSVQODIFLFPKVEKTKTPYIQVDIGGGQTK 120  
DB 61 AAKNOLTSNPENTVFDAKRLIGRTWNDSVQODIFLFPKVEKTKTPYIQVDIGGGQTK 120

Qy	121	TFABEELISAWLVTMKETAALVLTGKVVYNAVVTVPVAFNDADQROATKADGTAGLNVARI	180
Db	121	TFABEELISAWLVTMKETAALVLTGKVVYNAVVTVPVAFNDADQROATKADGTAGLNVARI	180
Qy	181	INEPTAAAIAYGLDQREGKENILVFDLGGSTPVSILITDNGFEVAVATNGTHLGGEDF	240
Db	181	INEPTAAAIAYGLDQREGKENILVFDLGGSTPVSILITDNGFEVAVATNGTHLGGEDF	240
Qy	241	DQVMEHFILKLYKKKTGKDVKRDKNRAVQKLRREVEKAKRALSSQOARIETIESFEYGEDF	300
Db	241	DQVMEHFILKLYKKKTGKDVKRDKNRAVQKLRREVEKAKRALSSQOARIETIESFEYGEDF	300
Qy	301	SETLTRAKFELNMDLFRSTMKPVQKLVLESDLSKSDIDEIYLVGGSTPIPKIQOLWEEF	360
Db	301	SETLTRAKFELNMDLFRSTMKPVQKLVLESDLSKSDIDEIYLVGGSTPIPKIQOLWEEF	360
Qy	361	FNGKEPSRGIPDPAVAVYGAAGVYSGSDQDYGDLVLDDVCPLTGTFYGGVMTKLIP	420
Db	361	FNGKEPSRGIPDPAVAVYGAAGVYSGSDQDYGDLVLDDVCPLTGTFYGGVMTKLIP	420
Qy	421	RNTVPTKKSQIFSTASDNOPTVYIKVYEGEERPLTDMHLGTFDTJGIPPARPGVQIE	480
Db	421	RNTVPTKKSQIFSTASDNOPTVYIKVYEGEERPLTDMHLGTFDTJGIPPARPGVQIE	480
Qy	481	VTFEIVDNGILRYVAEDKGGKGNKKTITITDQNRLLPEELERKRVNDAEKFAEDDKLKER	540
Db	481	VTFEIVDNGILRYVAEDKGGKGNKKTITITDQNRLLPEELERKRVNDAEKFAEDDKLKER	540
Qy	541	IDTNELESYAVSLKNOIGKREKLGKLSSEDKETMEKAVEEKEIEMLESODADIEDFKA	600
Db	541	IDTNELESYAVSLKNOIGKREKLGKLSSEDKETMEKAVEEKEIEMLESODADIEDFKA	600
Qy	601	KKKELEIVQPIISKLYGSAGPPPTGEEDTAEL	633
Db	601	KKKELEIVQPIISKLYGSAGPPPTGEEDTAEL	633
RESULT	2		
Q9NPF1	Q9NPF1	PRELIMINARY; PRT; 654 AA.	
AC	Q9NPF1		
DT	01-OCT-2000 (TREMblrel. 15, Created)		
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)		
DE	GLUCOSE-REGULATED PROTEIN (EMDOPLASMIC RETICULUM LUMENAL CA2+ BINDING PROTEIN GRP78).		
DE	HSPA5.		
GN	Homo sapiens (Human).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eulheta; Primates; Catarrhini; Homidae; Homo.		
OC	NCBI_taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=FIROBLASTS;		
RT	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.		
RT	"GRP78 is involved in the quality control of the LDL-receptor."		
RL	Submitted (JAN-2000) to the EMBL/Genbank/DBD databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=FIROBLASTS;		
RA	Chao C.;		
RL	Submitted (FEB-2000) to the EMBL/Genbank/DBD databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Bermudez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;		
RT	"Sequence differences between human grp78/BIP isolated from Hela cells		
RT	and previously reported human sequences."		
RL	Submitted (DEC-1999) to the EMBL/Genbank/DBD databases.		
RC	EMBL, AJ271729; CAB71335.1; -		
DR	EMBL, AF216292; AAF42836.1; -		
DR	InterPro: IPR000086; ER_Latarget.		
DR	InterPro: IPR001023; HSP70.		

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:57 ; Search time 18.76 Seconds  
(without alignments)  
1248.871 Million cell updates/sec

Title: US-09-806-955-1

Perfect score: 3273  
Sequence: 1 MEEDKREVDGTVGVGIDLGTT.....AGPPPTGEEDTAELHHHHHH 639

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

1 number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3216	98.3	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	98.0	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	97.9	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	97.8	654	1 GR78_RAT	P06761 rattus norv
5	3168	96.8	652	1 GR78_CHICK	090593 gallus gall
6	3069	93.8	658	1 GR78_XENLA	091883 xenopus lae
7	2722.5	83.2	667	1 GR78_APLCA	016556 aplysia cal
8	2708	82.7	656	1 HS7C_DROME	P29844 drosophila
9	2594.5	79.3	661	1 HS7C_CAMEL	P27420 caenorabd1
10	2356.5	72.0	667	1 BIP4_TOBAC	003684 nicotiana t
11	2347.5	71.7	668	1 BIP5_TOBAC	003685 nicotiana t
12	2332.5	71.3	663	1 BIP3_MAIZE	024581 zea mays (m
13	2328.5	71.1	666	1 BIP_LYCES	049118 lycopersico
14	2328.5	71.1	668	1 BIP_SPIOI	042334 splinacia ol
15	2322.5	71.0	663	1 BIP2_MAIZE	P24067 zea mays (m
16	2301	70.3	669	1 BIP1_ARATH	091KX3 arabidopsis
17	2299	70.2	668	1 BIP2_ARATH	039043 arabidopsis
18	2227.5	68.1	662	1 GR78_NEUCR	P78695 neuropept
19	2216	67.7	663	1 GR78_SCHPO	P36604 schizosacch
20	2195	67.0	670	1 GR78_YARLI	099170 yarrowia li
21	2174	66.4	682	1 GR78_YEAST	P16474 saccharomyc
22	2100	64.2	655	1 GR78_PLAFO	005866 plasmodium
23	2085.5	63.7	641	1 HS71_YEAST	P11051 saccharomyc
24	2083	63.6	646	1 HS7C_HUMAN	P11142 homo sapien
25	2083	63.6	646	1 HS7C_MOUSE	P08109 mus musculu
26	2079	63.5	650	1 HS7C_BOVIN	P19120 bos taurus
27	2078	63.5	644	1 HS71_PICAN	P53421 pichia angu
28	2076.5	63.4	640	1 HS7C_DICDI	P36415 dicystostell
29	2074.5	63.4	645	1 HS71_CANAL	P41797 cricetulus
30	2074	63.3	646	1 HS7C_CRIGR	P19378 cricetulus
31	2072	63.3	638	1 HS72_YEAST	P10592 saccharomyc
32	2065.5	63.1	679	1 GR78_KIULA	P22010 kluyveromyc
33	2065	63.1	648	1 HS71_PUCGR	Q01877 puccinia gr

34	2064	63.1	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	63.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.0	646	1 HS70_NEUCR	001233 neuropept
37	2060	62.9	643	1 HS71_SCHPO	010265 schizosacch
38	2059	62.9	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	62.7	651	1 HS7C_PETTHY	P09189 petunia hyb
40	2052	62.7	652	1 HS7D_MANSE	Q9u639 manduca sex
41	2049	62.6	651	1 HS70_ONCMY	P08108 oncorhynch
42	2046	62.5	641	1 HS72_BOVIN	Q27965 bos taurus
43	2045	62.5	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	62.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	62.3	641	1 HS71_HUMAN	P08107 homo sapien

## ALIGNMENTS

RESULT 1	ID	GR78_HUMAN	STANDARD:	PRT:	654 AA.
AC	P11021	09NPF1:			
DT	01-JUL-1989	(Rel. 11, Created)			
DT	20-AUG-2001	(Rel. 40, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN				
DE	HEAVY CHAIN BINDING PROTEIN) (BIP) (ENDOPLASMIC RETICULUM LUMENAL CA2+				
DE	BINDING PROTEIN GRP78).				
GN	HSP45 OR GRP78				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88283347; PubMed=2840249;				
RA	Ting J., Lee A.S.;				
RT	"Human gene encoding the 78,000-dalton glucose-regulated protein and				
RT	its pseudogene: structure, conservation, and regulation.";				
RL	DNA 7:275-286(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Cervix carcinoma;				
RA	Chao C.C.K.;				
RT	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fibroblast;				
RA	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;				
RT	"Grp78 is involved in the quality control of the LDL-receptor.";				
RT	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Bermudez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;				
RT	"Sequence differences between human grp78/BIP isolated from HeLa cells				
RT	and previously reported human sequences.";				
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 1-25 FROM N.A.				
RX	MEDLINE=93126087; PubMed=1480470;				
RA	Chao C.C.K., Lin-Chao S.;				
RT	"A direct-repeat sequence of the human BIP gene is required for				
RT	A23187-mediated inducibility and an inducible nuclear factor				
RT	binding.";				
RL	Nucleic Acids Res. 20:6481-6485(1992).				
RN	[6]				
RP	SEQUENCE OF 22-38.				
RC	TISSUE=Breast carcinoma;				
RX	MEDLINE=97295304; PubMed=9150946;				
RA	Resmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.;				
RT	"Two-dimensional electrophoretic analysis of human breast carcinoma				
RT	proteins: mapping of proteins that bind to the SH3 domain of mixed				
RT	lineage kinase MLK2.";				

```

RL Electrophoresis 18:588-598(1997).
RN (7)
RP SEQUENCE OF 19-40.
RC TISSUE-Colon carcinoma;
RX MEDLINE-97295306; PubMed-9150948;
RA JI H., Reid C.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins."
CC Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: M19645; AAA52614.1;
DR EMBL: X87949; CA61201.1;
DR EMBL: AJ271729; CAB71335.1;
DR EMBL: AF216292; AAF42836.1;
DR EMBL: X59969; CAA42595.1;
DR PIR: A29821; A29821.
DR HSSP: P19120; 1ATR.
DR SWISS-2DPAGE: P11021; HUMAN.
DR MIM: 138120;
DR InterPro: IPR000886; ER-target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER-TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR KX ATP-binding; Endoplasmic reticulum; signal.
DR FT SIGNAL 1 18
DR FT CHAIN 19 654
DR FT SITE 651 654 PREVENT SECRETION FROM ER.
DR FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
DR FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
DR FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
DR FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
DR SQ SEQUENCE 654 AA: 72333 MW: 5987085832A00 CRC64;

1 Query Match 98.3%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 1.5e-151;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEDKEDVGVGIDLTGTYSCVGFKNRVEIANDGCRITPSYAFTEGGERLIGA 61
DB 20 EEDKEDVGVGIDLTGTYSCVGFKNRVEIANDGCRITPSYAFTEGGERLIGA 79
QY 62 AKNQLTSNPENTVPDAKRLIGRTWNPVSQODIKFLPEVVEKKRKYQVNDIGGQRT 121
DB 80 AKNQLTSNPENTVPDAKRLIGRTWNPVSQODIKFLPEVVEKKRKYQVNDIGGQRT 139
QY 122 FAPETISAVVLTMKETAAVYLGKVTAVVPAVFNDQAQATKAGTIGLVNMRIT 181
DB 140 FAPETISAVVLTMKETAAVYLGKVTAVVPAVFNDQAQATKAGTIGLVNMRIT 199
QY 182 NEPTAAAIAYGLDKRGEENILVFDLGGSTFVSLITDNGVEVAVATNGDTHLGGEDFD 241
DB 200 NEPTAAAIAYGLDKRGEENILVFDLGGSTFVSLITDNGVEVAVATNGDTHLGGEDFD 259
QY 242 QVMEHFITLYKKKTKGKDVKNRAVQKLRREVEKAKRALSSQHARITESEFSEDEDS 301
DB 260 QVMEHFITLYKKKTKGKDVKNRAVQKLRREVEKAKRALSSQHARITESEFSEDEDS 319

```

```

QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSIDELIVLGGSTRIPKIQOLVKEFF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSIDELIVLGGSTRIPKIQOLVKEFF 379
QY 362 NKEKPSRGINPDEAVAYGAAGVLSQDPTGDIYLDVCPITLTGIEVGVWTKLIR 421
DB 380 NKEKPSRGINPDEAVAYGAAGVLSQDPTGDIYLDVCPITLTGIEVGVWTKLIR 439
QY 422 NVVPPKKQISFSTASDNPOTVITIKVEGERPLTKDNHLGTFEDLTGIPARGVQPIEV 481
DB 440 NVVPPKKQISFSTASDNPOTVITIKVEGERPLTKDNHLGTFEDLTGIPARGVQPIEV 499
QY 482 TFEIDVNGILRYAEDKGTGNKNTITTDQNRLPPEELIERYVNDAEKFAEDDKLKERI 541
DB 500 TFEIDVNGILRYAEDKGTGNKNTITTDQNRLPPEELIERYVNDAEKFAEDDKLKERI 559
QY 542 DTRNELESYAYSLKNDIGKEKLGKLSSEDEKTEMEKAVEEIEWLSHQDDIEDFKAK 601
DB 560 DTRNELESYAYSLKNDIGKEKLGKLSSEDEKTEMEKAVEEIEWLSHQDDIEDFKAK 619
QY 602 KKELEIVQPIISKLYGSAGPPTGEEDPAE 632
DB 620 KKELEIVQPIISKLYGSAGPPTGEEDPAE 650

RESULT 2
GR78_MESAU STANDARD; PRT; 654 AA.
ID GR78_MESAU
AC P07823;
DT 01-AUG-1988 (rel. 08, Created)
DT 01-AUG-1988 (rel. 08, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE 78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN
DE HEAVY CHAIN BINDING PROTEIN) (BIP).
GN HSPAS OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-87305586; PubMed-3623104;
RX Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat."
RL Gene 55:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: M17169; AAA51448.1;
DR PIR: A27414; A27414.
DR HSSP: P19120; 1ATR.
DR InterPro: IPR000886; ER-target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER-TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR KX ATP-binding; Endoplasmic reticulum; signal.

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 30.91 Seconds

(Without alignments)  
1574.751 Million cell updates/sec

Title: US-09-806-955-1

Perfect score: 3273

Sequence: 1 MEDKKEDVGTGVDIGLGT.....AGPPTGEGDTAELHHHHH 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	2 A27414	dnak-type molecule
2	3201	97.8	655	2 A37048	dnak-type molecule
3	3200	97.8	654	1 HHPRGB	dnak-type molecule
4	3182.5	97.2	653	2 A29821	dnak-type molecule
5	3168	96.8	652	2 I50242	dnak-type molecule
6	2722.5	83.2	667	2 S24782	dnak-type molecule
7	2719.5	83.1	667	2 D44261	dnak-type molecule
8	2710	82.8	656	2 JN0666	heat shock 70k pro
9	2618.5	80.0	661	2 T15513	dnak-type molecule
10	2594.5	79.3	661	2 A32475	dnak-type molecule
11	2559	78.2	657	2 T34037	heat shock 70k pro
12	2356.5	72.0	667	2 S21879	dnak-type molecule
13	2347.5	71.7	668	2 S21880	dnak-type molecule
14	2332.5	71.3	663	2 T04080	dnak-type molecule
15	2330	71.2	668	2 T46574	dnak-type molecule
16	2322.5	71.0	663	2 T04078	dnak-type molecule
17	2305.5	70.4	663	2 T03581	dnak-type molecule
18	2298	70.2	668	2 S71171	dnak-type molecule
19	2285	69.8	672	2 T43723	dnak-type molecule
20	2278	69.6	672	2 T43716	dnak-type molecule
21	2261	69.1	642	2 T45298	dnak-type molecule
22	2240.5	68.5	654	2 T06598	dnak-type molecule
23	2237	68.3	655	2 H86222	hypothetical prote
24	2231	68.1	658	2 S38890	dnak-type molecule
25	2227.5	68.1	662	2 T50464	glucose-regulated
26	2216	67.7	663	2 T38155	78 kd glucose regu
27	2214.5	67.4	608	2 T05741	dnak-type molecule
28	2205	66.4	666	2 T06358	dnak-type molecule
29	2174	66.4	682	1 HHBYK2	dnak-type molecule

## ALIGNMENTS

RESULT	1	ALIGNMENTS
A27414	dnak-type molecular chaperone GRP78 precursor - Chinese hamster	
N/Alternate names: glucose-regulated 78k protein; Ig heavy chain-binding protein		
C/Species: Citicellus griseus (Chinese hamster)		
C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999		
C/Accession: A27414		
R/Ref: J. Wooden, S.K.; Kriz, R.; Kelleher, K.; Kautman, R.J.; Lee, A.S.		
Gene 55, 147-152, 1987		
A/Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protein		
A/Reference number: A27414; M0ID:87305586		
A/Accession: A27414		
A/Molecule type: mRNA		
A/Residues: 1-654 <PIN>		
A/Cross-references: GB:M17169; NID:q191090; PIRN:AAA51448.1; PID:q304510		
C/Genetics:		
A/Gene: GRP78		
C/Function:		
A/Description: Involved in protein folding and assembling/dissassembling of protein co		
C/Superfamily: heat shock protein 70		
C/Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone		
F/1-18/Domain: signal sequence #status predicted <Sig>		
F/19-654/Product: Ig heavy chain binding protein #status predicted <MAT>		
F/651-654/Region: endoplasmic reticulum retention signal		
Query Match	98.0%	Score 3209; DB 2; Length 654;
Best Local Similarity	99.7%	Pred. No. 7.3e-155;
Matches 629; Conservative	2; Mismatches 0; Indels 0; Gaps 0;	
Oy	2 EEDKKEDVGTGVDIGLGTTCSCVGFKNRGVEIIANDQGNRITPSYAFPEGRLGDA 61	dnak-type molecule
Db	20 EEDKKEDVGTGVDIGLGTTCSCVGFKNRGVEIIANDQGNRITPSYAFPEGRLGDA 79	dnak-type molecule
Oy	62 AKNOLTSNPENTVDARKLIGRTWNDSVOODIKFLPFKVEKTKRYIOVDIGGGQTKT 121	dnak-type molecule
Db	80 AKNOLTSNPENTVDARKLIGRTWNDSVOODIKFLPFKVEKTKRYIOVDIGGGQTKT 139	dnak-type molecule
Oy	122 FAPREISAMVLTTKKETAELVGLKKTTHAVTPVAYVNDQROTKRAGTACLNWRII 181	dnak-type molecule
Db	140 FAPREISAMVLTTKKETAELVGLKKTTHAVTPVAYVNDQROTKRAGTACLNWRII 199	dnak-type molecule
Oy	182 NEPTAAAIAGLDKRESEKNILVFDLGGTFVSLITINDGVFEVAVTNGDTHLGGEDFD 241	dnak-type molecule
Db	200 NEPTAAAIAGLDKRESEKNILVFDLGGTFVSLITINDGVFEVAVTNGDTHLGGEDFD 259	dnak-type molecule
Oy	242 ORVNEHFTIKLYKKTKGDKVRKNRAVQKLREVEKARALSSQHOARIEIESPEGEDFS 301	dnak-type molecule
Db	260 ORVNEHFTIKLYKKTKGDKVRKNRAVQKLREVEKARALSSQHOARIEIESPEGEDFS 319	dnak-type molecule
Oy	302 ETLLRAKFEELNMLFSTKMPVKVLEDSDLKSDIDEIVLGGSRIRIKIQLVNEFF 361	dnak-type molecule

Db 320 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLGVSGSTRIPKIQOLVKEPF 379  
 QY 362 NKEPERGINPDEAVAYGAAGVLSGDDPGLDVLVPCPLTIGIEVGVMTKLIR 421  
 Db 380 NKEPERGINPDEAVAYGAAGVLSGDDPGLDVLVPCPLTIGIEVGVMTKLIR 439  
 QY 422 NTVPPTKKSQIFSTASDNDPTVITIKYEGEERPLTKDNHLLGTFDLGIPPARGVQIEV 481  
 Db 440 NTVPPTKKSQIFSTASDNDPTVITIKYEGEERPLTKDNHLLGTFDLGIPPARGVQIEV 499  
 QY 482 TFEIDVNGILRVTAEDKGTKGNKKITITNDQNRLPPEEIERMVNDAEKFAEEDKKLKERI 541  
 Db 500 TFEIDVNGILRVTAEDKGTKGNKKITITNDQNRLPPEEIERMVNDAEKFAEEDKKLKERI 559  
 QY 542 DTRNELESTAYSLKNOIGKEKLGKLSSEDEKTMKAEEKTEWLESHQDADIEDFKAK 601  
 Db 560 DTRNELESTAYSLKNOIGKEKLGKLSSEDEKTMKAEEKTEWLESHQDADIEDFKAK 619  
 QY 602 KKELEIVOPITISKLYGSAGPPTGEEDYAE 632  
 Db 620 KKELEIVOPITISKLYGSAGPPTGEEDYAE 650

## RESULT 2

A37048

dnak-type molecular chaperone grp78 precursor - mouse

N.Alternate names: Bfp: glucose-regulated 78k protein; Ig heavy chain-binding protein

C.Species: Mus musculus (house mouse)

C.Date: 15-Jan-1993 #sequence-revision 15-Jan-1993 #text-change 20-Aug-1999

C.Accession: A37048; B37048; J00094; I49137; A31934

R.Kozlitski, Y.; Northington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.

J. Cell Sci. Suppl. 11, 115-137, 1989

A.Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulat

A.Reference number: A37048; MUID:90130686

A.Accession: A37048

A.Status: not compared with conceptual translation

A.Molecule type: mRNA

A.Residues: 1-655 &lt;K21&gt;

A.Accession: B37048

A.Molecule type: protein

A.Residues: 20-41, 'X', 43-46 &lt;K22&gt;

R.Paretti, C.L.O.; Hofbauer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.

Gene 82, 291-303, 1989

A.Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero

A.Reference number: J00094; MUID:90060818

A.Accession: J00094

A.Molecule type: mRNA

A.Residues: 488-590, 'G', 592-595, 'E', 597-655 &lt;PAR&gt;

A.Cross-references: GB:M30779; NID:9193644; PIDN:AAA37742.1; PID:9193645

R.Tillman, J.B.; Mote, P.L.; Walford, R.L.; Spindler, S.R.

Gene 158, 225-229, 1995

A.Title: Structure and regulation of the mouse GRP78 (BiP) promoter by glucose and calci

A.Reference number: I49137; MUID:95331621

A.Accession: I49137

A.Status: translation not shown; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-28 &lt;RES&gt;

A.Cross-references: EMBL:U16277; NID:9829364; PIDN:AAA76734.1; PID:9829365

R.Haas, I.G.; Moe, T.

Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988

A.Title: cDNA cloning of the immunoglobulin heavy chain binding protein.

A.Reference number: A31934; MUID:88176922

A.Accession: A31934

A.Molecule type: mRNA

A.Residues: 514-589, 'D', 591-595, 'E', 597-655 &lt;HAA&gt;

A.Cross-references: GB:M19351; NID:9192231; PIDN:AAA37315.1; PID:9387113

C.Genetics:

A.Gene: grp78

C.Function:

Involved in protein folding and assembling/disassembling of protein comp

C.Superfamily: heat shock protein 70

C.Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon

F.1-19/Domains: signal sequence #status predicted &lt;SIC&gt;

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>  
 F:652-655/Region: endoplasmic reticulum retention signal

Query Match 97.8%; Score 3201; DB 2; Length 655;  
 Best Local Similarity 99.4%; Pred. No. 1,8e-154;  
 Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EEDKKEVDVGVGIDGTTSCVGVFKNKRVELIANDQNRITPSPVATPGEERLIGA 61  
 Db 21 EEDKKEVDVGVGIDGTTSCVGVFKNKRVELIANDQNRITPSPVATPGEERLIGA 80  
 QY 62 AKNQLTSNENPVFPDAKRLIGRTWNPVSQODIKFLPFVVEKTKTPYQVNDIGGQRT 121  
 Db 81 AKNQLTSNENPVFPDAKRLIGRTWNPVSQODIKFLPFVVEKTKTPYQVNDIGGQRT 140  
 QY 122 FAPPEISAMVLTQKMETAEAYLGKVVTHAVVTPAYFNDQROATKDACTAGLANVRI 181  
 Db 141 FAPPEISAMVLTQKMETAEAYLGKVVTHAVVTPAYFNDQROATKDACTAGLANVRI 200  
 QY 182 NEPTAAIAVGLDKRGEKNILVFDLGGCTPPVSLITDNGVFEVATNGDTHLGGEDPD 241  
 Db 201 NEPTAAIAVGLDKRGEKNILVFDLGGCTPPVSLITDNGVFEVATNGDTHLGGEDPD 260  
 QY 242 QRYMEHFILYKTKTKGDKVDRKDNRAVQKLRREYKAKRALSSQHQARIESFEYGEDPS 301  
 Db 261 QRYMEHFILYKTKTKGDKVDRKDNRAVQKLRREYKAKRALSSQHQARIESFEYGEDPS 320  
 QY 302 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLGVSGSTRIPKIQOLVKEPF 361  
 Db 321 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLGVSGSTRIPKIQOLVKEPF 380  
 QY 362 NKEPERGINPDEAVAYGAAGVLSGDDPGLDVLVPCPLTIGIEVGVMTKLIR 421  
 Db 381 NKEPERGINPDEAVAYGAAGVLSGDDPGLDVLVPCPLTIGIEVGVMTKLIR 440  
 QY 422 NTVPPTKKSQIFSTASDNDPTVITIKYEGEERPLTKDNHLLGTFDLGIPPARGVQIEV 481  
 Db 441 NTVPPTKKSQIFSTASDNDPTVITIKYEGEERPLTKDNHLLGTFDLGIPPARGVQIEV 500  
 QY 482 TFEIDVNGILRVTAEDKGTKGNKKITITNDQNRLPPEEIERMVNDAEKFAEEDKKLKERI 541  
 Db 501 TFEIDVNGILRVTAEDKGTKGNKKITITNDQNRLPPEEIERMVNDAEKFAEEDKKLKERI 560  
 QY 542 DTRNELESTAYSLKNOIGKEKLGKLSSEDEKTMKAEEKTEWLESHQDADIEDFKAK 601  
 Db 561 DTRNELESTAYSLKNOIGKEKLGKLSSEDEKTMKAEEKTEWLESHQDADIEDFKAK 620  
 QY 602 KKELEIVOPITISKLYGSAGPPTGEEDYAE 632  
 Db 621 KKELEIVOPITISKLYGSAGPPTGEEDYAE 651

## RESULT 3

HHRTGB

dnak-type molecular chaperone precursor - rat

N.Alternate names: Bfp: glucose-regulated 78k protein; Ig heavy chain-binding protein

N.Names: stereoidogenesis-activator polypeptide

C.Species: Rattus norvegicus (Norway rat)

C.Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 22-Jun-1999

C.Accession: A23948; A60134; A26257

R.Munro, S.; Pelham, H.R.B.

Cell 46, 291-300, 1986

A.Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p

A.Reference number: A23948; MUID:86245075

A.Accession: A23948

A.Molecule type: mRNA

A.Residues: 1-654 &lt;MUN&gt;

A.Cross-references: GB:M14050; NID:9203150; PIDN:AAA40817.1; PID:9203151

R.Pedersen, R.C.; Browlie, A.C.

Science 236, 188-190, 1987

A.Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor.

A.Reference number: A60134; MUID:87177981

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 43 Seconds

(without alignments)  
1100.764 Million cell updates/sec

Title: US-09-806-955-1

Perfect score: 3273  
Sequence: 1 MEEDKKEDVGVVIGIDICTT.....AGPPTGEEDTAELHHNNHH 639

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

al number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDSR/gcgdata/geneseq/AA1980.DAT.\*  
2: /SIDSR/gcgdata/geneseq/AA1981.DAT.\*  
3: /SIDSR/gcgdata/geneseq/AA1982.DAT.\*  
4: /SIDSR/gcgdata/geneseq/AA1983.DAT.\*  
5: /SIDSR/gcgdata/geneseq/AA1984.DAT.\*  
6: /SIDSR/gcgdata/geneseq/AA1985.DAT.\*  
7: /SIDSR/gcgdata/geneseq/AA1986.DAT.\*  
8: /SIDSR/gcgdata/geneseq/AA1987.DAT.\*  
9: /SIDSR/gcgdata/geneseq/AA1988.DAT.\*  
10: /SIDSR/gcgdata/geneseq/AA1989.DAT.\*  
11: /SIDSR/gcgdata/geneseq/AA1990.DAT.\*  
12: /SIDSR/gcgdata/geneseq/AA1991.DAT.\*  
13: /SIDSR/gcgdata/geneseq/AA1992.DAT.\*  
14: /SIDSR/gcgdata/geneseq/AA1993.DAT.\*  
15: /SIDSR/gcgdata/geneseq/AA1994.DAT.\*  
16: /SIDSR/gcgdata/geneseq/AA1995.DAT.\*  
17: /SIDSR/gcgdata/geneseq/AA1996.DAT.\*  
18: /SIDSR/gcgdata/geneseq/AA1997.DAT.\*  
19: /SIDSR/gcgdata/geneseq/AA1998.DAT.\*  
20: /SIDSR/gcgdata/geneseq/AA1999.DAT.\*  
21: /SIDSR/gcgdata/geneseq/AA2000.DAT.\*  
22: /SIDSR/gcgdata/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	21	AAV90693
2	3225	98.5	633	21	AAV90694
3	3209	98.0	654	15	AAV53075
4	3182.5	97.2	653	22	AAH82533
5	3171	96.9	656	14	AAH34934
6	3168	96.8	666	15	AAH33076
7	2299	70.2	668	21	AAH41382
8	2299	70.2	691	21	AAH41381
9	2258.5	69.0	642	21	AAH39285
10	2258.5	69.0	658	21	AAH39284
11	2237	68.3	655	21	AAH30912
					Arabidopsis thalia

12	2174	66.4	682	22	AAH70749	S cerevisiae apoct
13	2125.5	64.9	687	22	AAH70875	C albicans apoptos
14	2083	63.6	646	19	AAH54364	Heat shock cognate
15	2083	63.6	646	20	AAH17407	Human heat shock c
16	2083	63.6	646	20	AAH17408	Mouse heat shock c
17	2083	63.6	646	21	AAH23649	Mouse heat shock p
18	2083	63.6	646	21	AAH23651	Human heat shock p
19	2083	63.6	646	22	AAH22535	Human heat shock p
20	2083	63.6	890	21	AAH22938	GFP-HSC70 fusion p
21	2074.5	63.4	656	18	AAH01638	Candida albicans h
22	2064	63.1	646	11	AAH03927	Rat HSP (ratHSP70)
23	2041	62.4	647	11	AAH03928	Xenopus laevis HSP
24	2039.5	62.3	624	21	AAH23252	Human Hsp72 (heat
25	2039.5	62.3	641	19	AAH54349	Human heat shock 7
26	2039.5	62.3	641	21	AAH33652	Human heat shock p
27	2039.5	62.3	641	22	AAH82534	Human heat shock p
28	2036.5	62.2	646	21	AAH53604	Arabidopsis thalia
29	2033.5	62.1	633	14	AAH43002	Mouse Slp1p homolo
30	2032.5	62.1	647	20	AAH1380	T. gondii antigen
31	2032.5	62.1	647	22	AAH49099	Toxoplasma gondii
32	2032	62.1	634	11	AAH03930	Gallus gallus HSP
33	2032	61.7	640	21	AAH23653	Human heat shock p
34	2017.5	61.6	640	11	AAH03929	Homo sapiens HSP (
35	2014	61.5	642	21	AAH23650	Rat heat shock pro
36	2013	61.5	640	18	AAH10065	Human heat shock p
37	2013	61.5	640	21	AAH8408	Human heat shock p
38	2013	61.5	677	11	AAH09418	Hsp70 antigen from
39	2008	61.4	641	14	AAH43004	Mouse sperm
40	2004.5	61.2	669	11	AAH03925	T. cruzi HSP (tcr70
41	1994.5	60.9	676	9	AAH80088	Sequence of 70kd p
42	1974	60.3	665	21	AAH58386	Lung cancer associ
43	1970	60.2	646	11	AAH03931	Zea mays HSP (mzh
44	1968.5	60.1	657	11	AAH03926	T. cruzi HSP. Try
45	1963.5	60.0	623	21	AAH53605	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
ID AAV90693 standard; Protein: 639 AA.  
XX AAV90693;  
AC  
XX 29-AUG-2000 (first entry)  
DT  
XX 639 residue human immunoglobulin heavy binding protein, BIP(78KD).  
DE  
XX Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;  
KW autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.  
KM  
XX  
OS Homo sapiens.  
XX  
XX WO200021995-A1.  
XX  
XX 20-Apr-2000.  
PD  
XX  
XX 08-OCT-1999; 99WO-GB03316.  
PF  
XX  
XX 09-OCT-1998; 98GB-0022115.  
PR  
XX  
XX (UNLO ) KINGS COLLEGE LONDON.  
PA  
XX  
XX Panayl GS, Corriganl VM, Bodman-Smith MD, Fife MS, Lanchbury JS;  
FI WPI: 2000-317942/27.  
XX N-PSDB: AAA30792.  
DR  
XX New human immunoglobulin heavy chain binding protein and encoding  
PT polynucleotide, useful for diagnosis and treatment of rheumatoid  
arthritis -  
XX

PS Claim 3; Page 23; 53pp; English.

This sequence represents a human immunoglobulin heavy chain binding protein, B1P(78KD), having a 639 amino acid sequence. The invention also encompasses a B1P(78KD) protein of 633 amino acids (Y90694). The cDNA encoding B1P(78KD), also referred to as p78 in the specification, was isolated from human chondrocytes (the specialised cells of articular cartilage) and human chondrosarcoma cell lines. The B1P(78KD) cDNA of this invention contains a number of differences compared with the published sequence (Genbank accession number X87949) and has therefore been deposited with Genbank with the accession number AF186511. These differences comprise 6 single nucleotide substitutions and a codon insertion, and result in three amino acid substitutions and an arginine insertion at position 834-836 of the protein. The B1P(78KD) proteins react with antibodies present in the serum of rheumatoid arthritis patients, and is therefore a putative autoantigen for this autoimmune disease. B1P(78KD) is also able to selectively proliferate synovial T-cells from patients with rheumatoid arthritis. B1P(78KD) or peptides derived from the protein are useful reagents to indicate the presence of rheumatoid arthritis, and can be used in prognostic or diagnostic tests of body fluids for rheumatoid arthritis by ELISA (enzyme linked immunosorbent assay) or Western blotting. The protein or the cDNA encoding it can also be used to test for rheumatoid arthritis by detecting antibodies to the protein. B1P(78KD), its peptides and polynucleotides are also useful therapeutically.

**SQ**      **Sequence**      **639 AA;**

Query Match	100.0%	Score 3273;	DB 21;	Length 639;
Best Local Similarly	100.0%;	Pred. NO. 1.3e-209;		
Matches 639; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MEBKKKDDVGVVNOIDIGTTSYSCYGVKKNRGVEIANDOSNRITTPSVATPEGGERIGD	60
Db	1	meedkcedvclvgvaidigctescvgyvfkngvleilandgnrltprsvatfpegerrlsgd	60
Qy	61	AAKQOLSPNENPVFADKRLRGRTMNPSVQOODKFLPFYVVEKKTPYUOVIGGORT	120
Db	61	aakqglcsnpentvfoakrllrgtwnbpsvgdtkflprfvevkkckpryiqvdi99qtk	120
Qy	121	TEAPEISAMVLTMKMETAEVLEKKTVAHVTVPAYENDAOQATKDACTIAGLVNMR	180
Db	121	tfapeeisamvltmkmetaevlgkvtthavtvpayfndaqgactdagtlaglvnmrl	180
Qy	181	INEPTAAIAVGLDKRGEKENILVFDLGSGTFPVSULTDNGYFEVYATNGDHLGDEP	240
b	181	ineptaaiavgldkrgekenilvfdlgsgtfdslltlingfeyvatngdchlggedt	240
Qy	241	DQWMEHFIRYKKKTKDNRKDNRAQOKLRREVEKAKRALSOSHOARIESFYGEPD	300
Db	241	dqrwemhfirlykktkdknrkdnravqkrlrrevkakrlsngqarlesfeygedt	300
Qy	301	SETULTRAKFEEELNDDLRSTMKPVQVYLEDSDLKSSDIDEIVLVGSGTRIPKIQOLVKEP	360
Db	301	setltrakeelndmlfrstmkpvqvyledsdlkssdideivlvgsttripkidlqvkef	360
Qy	361	FNGKEPBRGRTPDRAVAVYGAAVQGVLSGQODIGDVLVLYCPLTIGIEVGVGMKRLIP	420
Db	361	fngkepsbrgtpdpdraavygaavqgvlsqgdtdvlvlycpltlgievsgvmkrlip	420
Qy	421	RNTVVPKRSQIFSTADNDOPVTIVIKYEEGERPLTKONHLEGFEDTLGTPAPARGPOIE	480
Db	421	rntvvpkksqifstaadndpvtivikyegeerpltkonhlgfctdlgtppaprgpqlie	480
Qy	481	VTFEIDVNGILRTVAEDKGTKNNKRTITNDONRLRPEELIRVYNDAEKFAEEDKRLKER	540
Db	481	vtfeidvngilrtvaedkgtknkrtitndonrlrpeelirmvndaekfaeedkrlker	540
Qy	541	IDTRNEELSAVYSLKNOIGOKREKAGLSSEDKETMKAVBEKIBWLESQODDIEDFKA	600
Db	541	idtrnelesavyalskngidgkexlgyglssedketmekaveeklewlespqddiedfka	600

[illegible]

## RESULT 2

ID AAY90694 standard; Protein; 633 AA.

AC AAY90694;

DT 29-AUG-2000 (first entry)

DE 633 residue human immunoglobulin heavy binding protein, BiP(78KD).

KM Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;  
KM autoantigen: rheumatoid arthritis; antiarthritic; antirheumatic; p78

OS Homo sapiens.

PN W0200021995-A1.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

PA (UNLO ) KINGS COLLEGE LONDON.

PI Panayl GS, Corriganl VM, Bodman-Smith MD, Flfe MS, Lanchbury JS;

DR WPI; 2000-317942/27.

PT New human immunoglobulin heavy chain binding protein and encoding polynucleotide, useful for diagnosis and treatment of rheumatoid arthritis -

PS Claim 3; Page 44-46; 53pp; English.

CC This sequence represents a human immunoglobulin heavy chain binding  
CC protein, BiP(78kd), having a 633 amino acid sequence. The invention  
CC also encompasses a BiP(78kd) protein of 639 amino acids (Y906933).  
CC The cDNA encoding BiP(78kd), also referred to as p78 in the  
CC specification, was isolated from human chondrocytes (the specialised  
CC cells of articular cartilage) and human chondrosarcoma cell lines. The  
CC BiP(78kd) cDNA of this invention contains a number of differences  
CC compared with the published sequence (Genbank accession number X87949),  
CC and has therefore been deposited with Genbank with the accession  
CC number AF186811). These differences comprise 6 single nucleotide  
CC substitutions and a codon insertion, and result in three amino acid  
CC substitutions and an arginine insertion at position 834-836 of the  
CC protein. The BiP(78kd) proteins react with antibodies present in the  
CC serum of rheumatoid arthritis patients, and is therefore a putative  
CC autoantigen for this autoimmune disease. BiP(78kd) is also able to  
CC selectively proliferate synovial T-cells from patients with rheumatoid  
CC arthritis. BiP(78kd) or peptides derived from the protein are useful as  
CC reagents to indicate the presence of rheumatoid arthritis, and can be  
CC used in prognostic or diagnostic tests of body fluids for rheumatoid  
CC arthritis by ELISA (enzyme linked immunosorbent assay) or western  
CC blotting. The protein or the cDNA encoding it can also be used to test  
CC for rheumatoid arthritis by detecting antibodies to the protein.  
CC BiP(78kd), its peptides and polynucleotides are also useful  
CC therapeutically.

SD Sequence 633 AA;

Query Match	98.5%	Score 3225;	DB 21	Length 633;
Best Local Similarity	100.0%	Pred. No. 1.9e-206;		
Matches 633; Conservative	0;	Mismatches	0;	Gaps 0



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:37 ; Search time 50.77 Seconds

(Without alignments)  
1841.009 Million cell updates/sec

Title: US-09-806-955-1

Perfect score: 3273  
Sequence: 1 MEEDKEDVGTWVGIDLTGTT.....AGPPTGDEDTAELHHHHH 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3273	100.0	639	4 Q9UK02	Q9UK02 homo sapien
2	3216	98.3	654	4 Q9NPF1	Q9NPF1 homo sapien
3	3199	97.7	655	11 Q9DC41	Q9DC41 mus musculi
4	2994	91.5	658	13 Q91688	Q91688 xenopus lae
5	2725	83.3	658	5 Q76180	Q76180 bombyx mori
6	2717	83.0	656	5 Q9VYU3	Q9VYU3 drosophila
7	2607.5	79.7	656	5 Q62568	Q62568 suberilab d
8	2559	78.2	657	5 Q02085	Q02085 caenorhabdi
9	2534	77.4	651	5 Q24798	Q24798 echinococcu
10	2525.5	77.2	649	5 Q24895	Q24895 echinococcu
11	2460	75.2	648	5 Q45038	Q45038 schistosoma
12	2345.5	71.7	665	10 Q9M4E8	Q9M4E8 cucumis sat
13	2330	71.2	668	10 Q22639	Q22639 glycine max
14	2313.5	70.7	675	10 Q40924	Q40924 pseudotsuga
15	2305.5	70.4	663	10 Q24182	Q24182 oryza sativ
16	2292	70.0	668	10 Q9FSY7	Q9FSY7 corvus ave
17	2285	69.8	672	3 Q14453	Q14453 aspergillus
18	2278	69.6	672	3 Q13280	Q13280 aspergillus
19	2273	69.4	668	5 Q9U540	Q9U540 toxoplasma

20	2263	69.1	669	3 Q9C1C1	Q9C1C1 cronartium
21	2261	69.1	642	5 Q61001	Q61001 toxoplasma
22	2246.5	68.6	672	3 Q9UWE3	Q9UWE3 aspergillus
23	2243	68.5	659	10 Q41074	Q41074 phaeodactyl
24	2240.5	68.5	664	10 Q39830	Q39830 glycine max
25	2237	68.3	655	10 Q04022	Q04022 arabidopsis
26	2231	68.2	658	10 Q12752	Q12752 phytophthor
27	2229.5	68.1	701	5 Q24928	Q24928 elmeria ten
28	2214.5	67.7	608	10 Q40058	Q40058 hordeum vul
29	2205	67.4	666	10 Q39804	Q39804 glycine max
30	2188.5	66.9	655	3 P87028	P87028 pneumocysti
31	2160.5	66.0	665	3 Q9HG01	Q9HG01 pichia angu
32	2139.5	65.4	652	3 Q00053	Q00053 pneumocysti
33	2116.5	64.7	652	5 Q07615	Q07615 plasmodium
34	2108	64.4	651	3 Q9UVM1	Q9UVM1 rhizopus st
35	2093	63.9	659	5 Q9XJ22	Q9XJ22 crassostrea
36	2080	63.6	642	3 Q9UYM0	Q9UYM0 rhizopus st
37	2079	63.5	647	3 Q59855	Q59855 schizosacch
38	2076	63.4	661	5 Q77164	Q77164 entamoeba h
39	2074.5	63.4	648	10 Q40980	Q40980 pisum sativ
40	2069	63.2	646	13 Q73885	Q73885 gallus gall
41	2065.5	63.1	647	3 Q94104	Q94104 pneumocysti
42	2064.5	63.1	645	3 Q94106	Q94106 pneumocysti
43	2057	62.8	652	5 Q26924	Q26924 trypanosoma
44	2055.5	62.8	647	10 Q9M4E6	Q9M4E6 cucumis sat
45	2055.5	62.8	653	5 Q26733	Q26733 trypanosoma

## ALIGNMENTS

RESULT 1  
Q9UK02 PRELIMINARY; PRT; 639 AA.

AC Q9UK02;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE BIP PROTEIN (FRAGMENT).  
GN HSPA5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ARTICULAR CARTILAGE;  
RA File M.S.;  
RT \*BIP: an Autoantigen associated with Rheumatoid Arthritis.\*;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF186511; AAF13605.1; -  
DR HSSP: P19120; 3HSC.  
DR InterPro: IPR001023; HSP70.  
DR Pfam: PF00012; HSP70; 1.  
DR PRINTS: PR00301; HEATSHOCK70.  
DR PROSITE: PS00297; HSP70\_1; 1.  
DR PROSITE: PS00329; HSP70\_2; 1.  
DR PROSITE: PS01036; HSP70\_3; 1.  
FT NON\_TER 639  
SQ SEQUENCE 639 AA; 70931 MW; 3421208FP128FD5D CRC64;

Query Match 100.0%; Score 3273; DB 4; Length 639;  
Best Local Similarity 100.0%; Pred. No. 2.3e-158;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEDKEDVGTWVGIDLTGTVSGVGFKNRGRVETIANDGRRIPSYVAFTPEGERLIGD 60  
DB 1 MEEDKEDVGTWVGIDLTGTVSGVGFKNRGRVETIANDGRRIPSYVAFTPEGERLIGD 60  
QY 61 AAKNOLTSNPENTYFDKRLRGRTWNPVVOODIKFLPFKVEKKTPYIOVDIGGGQTK 120  
DB 61 AAKNOLTSNPENTYFDKRLRGRTWNPVVOODIKFLPFKVEKKTPYIOVDIGGGQTK 120

```

QY 121 TPAPEISAMVLTAKMKEAAYLGKKVTHAVVTPAYFNDQAQKATKAGTAGLNMRI 180
DB 121 TPAPEISAMVLTAKMKEAAYLGKKVTHAVVTPAYFNDQAQKATKAGTAGLNMRI 180
QY 181 INEPTAAAIAYGLDKREGKNIYFDLGGGTPVSLTTIDNGVEVAVATNGDTHLGGEDF 240
DB 181 INEPTAAAIAYGLDKREGKNIYFDLGGGTPVSLTTIDNGVEVAVATNGDTHLGGEDF 240
QY 241 DGRVMEHFILYKKTKGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYGEDF 300
DB 241 DGRVMEHFILYKKTKGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYGEDF 300
QY 301 SETLTRAKEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPIQOLYKKEF 360
DB 301 SETLTRAKEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPIQOLYKKEF 360
QY 361 FNGKEPSRGINPDEAVAYGAAGVAGVLSGDDOTGDLVLDVCPRLTGITGVGMKILIP 420
DB 361 FNGKEPSRGINPDEAVAYGAAGVAGVLSGDDOTGDLVLDVCPRLTGITGVGMKILIP 420
QY 421 RNTVPTKKSQIFSTASDNOPTVTIKYEGEERPLTKDNHLLGTFDLTGIPPARGVPOIE 480
DB 421 RNTVPTKKSQIFSTASDNOPTVTIKYEGEERPLTKDNHLLGTFDLTGIPPARGVPOIE 480
QY 481 VFPEIDVNGILRYTAEDKGTGNKNTITNDONRLTPPEIERKAVNDAEKFAEEDKKLKER 540
DB 481 VFPEIDVNGILRYTAEDKGTGNKNTITNDONRLTPPEIERKAVNDAEKFAEEDKKLKER 540
QY 541 IOTRNELESYAYSLKNOIGDKELGKLSSEDEKTEKAVEEKEIEMWLESHODADIEDFA 600
DB 541 IOTRNELESYAYSLKNOIGDKELGKLSSEDEKTEKAVEEKEIEMWLESHODADIEDFA 600
QY 601 KKELEEIYOPITISKLYGSAGPPPTGEEDTAELHHHHH 639
DB 601 KKELEEIYOPITISKLYGSAGPPPTGEEDTAELHHHHH 639

RESULT 2
Q9NPF1 PRELIMINARY: PRT: 654 AA.
AC 09NPF1:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GLUCOSE-REGULATED PROTEIN (ENDOPLASMIC RETICULUM LUMENAL CA2+ BINDING
DE PROTEIN GRP78).
GN HSP95.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=FIBROBLASTS;
RA Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;
RT "Grp78 is involved in the quality control of the IdL-receptor.";
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=FIBROBLASTS;
RA Chao C.;
RN 13
RP SEQUENCE FROM N.A.
RC TISSUE=FIBROBLASTS;
RA Bernudez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;
RT "Sequence differences between human grp78/BiP isolated from Hela cells
RT and previously reported human sequences.";
RN 14
RP EMBL; AJ271729; CAB71335.1; -
RN 15
RP EMBL; AF216293; AAF42836.1; -
RN 16
RP InterPro; IPR000886; ER_target.
RN 17
RP InterPro; IPR001023; HSP70.

```

```

DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS1036; HSP70_3; 1.
SQ SEQUENCE 654 AA; 72332 MW; 5987DBD85BC32A00 CRC64;

Query Match
Best Local Similarity 98.3%; Score 3216; DB 4; Length 654;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEDKKEVDGTVVIGIDLGTTVSCGVFKNGRVEIANDQGNRTTPSVATPGEERLIGDA 61
DB 20 EEDKKEVDGTVVIGIDLGTTVSCGVFKNGRVEIANDQGNRTTPSVATPGEERLIGDA 79
QY 62 AKNQLTSNPEENTVFDAKRLIGRTWNPVQODIKFLPFKVEKKRPYIQVDIGGQTKT 121
DB 80 AKNQLTSNPEENTVFDAKRLIGRTWNPVQODIKFLPFKVEKKRPYIQVDIGGQTKT 139
QY 122 PAPEISAMVLTAKMKEAAYLGKKVTHAVVTPAYFNDQAQKATKAGTAGLNMRI 181
DB 140 PAPEISAMVLTAKMKEAAYLGKKVTHAVVTPAYFNDQAQKATKAGTAGLNMRI 199
QY 182 NEPTAAAIAYGLDKREGKNIYFDLGGGTPVSLTTIDNGVEVAVATNGDTHLGGEDF 241
DB 200 NEPTAAAIAYGLDKREGKNIYFDLGGGTPVSLTTIDNGVEVAVATNGDTHLGGEDF 259
QY 242 ORVMEHFILYKKTKGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYGEDF 301
DB 260 ORVMEHFILYKKTKGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYGEDF 319
QY 302 ETLTRAKEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPIQOLYKKEF 361
DB 320 ETLTRAKEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPIQOLYKKEF 379
QY 362 NGKEPSRGINPDEAVAYGAAGVAGVLSGDDOTGDLVLDVCPRLTGITGVGMKILIP 421
DB 380 NGKEPSRGINPDEAVAYGAAGVAGVLSGDDOTGDLVLDVCPRLTGITGVGMKILIP 439
QY 422 RNTVPTKKSQIFSTASDNOPTVTIKYEGEERPLTKDNHLLGTFDLTGIPPARGVPOIE 481
DB 440 RNTVPTKKSQIFSTASDNOPTVTIKYEGEERPLTKDNHLLGTFDLTGIPPARGVPOIE 499
QY 482 VFPEIDVNGILRYTAEDKGTGNKNTITNDONRLTPPEIERKAVNDAEKFAEEDKKLKER 541
DB 500 VFPEIDVNGILRYTAEDKGTGNKNTITNDONRLTPPEIERKAVNDAEKFAEEDKKLKER 559
QY 542 IOTRNELESYAYSLKNOIGDKELGKLSSEDEKTEKAVEEKEIEMWLESHODADIEDFA 601
DB 560 IOTRNELESYAYSLKNOIGDKELGKLSSEDEKTEKAVEEKEIEMWLESHODADIEDFA 619
QY 602 KKELEEIYOPITISKLYGSAGPPPTGEEDTAELHHHHH 632
DB 620 KKELEEIYOPITISKLYGSAGPPPTGEEDTAELHHHHH 650

RESULT 3
Q9DC41 PRELIMINARY: PRT: 655 AA.
AC 09DC41:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).
GN HSP95.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:34 ; Search time 30.91 Seconds  
(without alignments)  
1559.964 Million cell updates/sec

Title: US-09-806-955-2  
Perfect score: 3225  
Sequence: 1 MEDKKEDVGVVIGLGT.....SKLYGSAGPPTEEDTAEI 633

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
a1 number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	99.5	654	2	A27414
2	3201	99.3	655	2	A37048
3	3200	99.2	654	1	HHR7GB
4	3182.5	98.7	653	2	A29821
5	3168	98.2	652	2	I50242
6	2722.5	84.4	667	2	S24782
7	2719.5	84.3	667	2	D44261
8	2710	84.0	656	2	JN0666
9	2618.5	81.2	661	2	T15513
10	2594.5	80.4	661	2	A34475
11	2555	79.2	657	2	T34037
12	2354.5	73.0	667	2	S21879
13	2346.5	72.8	668	2	S21880
14	2333.5	72.3	663	2	T04080
15	2325	72.1	668	2	T46574
16	2322.5	72.0	663	2	T04078
17	2305.5	71.5	663	2	T03581
18	2295	71.2	668	2	S71171
19	2283	70.8	672	2	T43723
20	2276	70.6	672	2	T43716
21	2261	70.1	642	2	T45298
22	2257	69.4	655	2	H86222
23	2234.5	69.3	664	2	T06598
24	2221	69.2	658	2	S38890
25	2227.5	69.1	662	2	T50464
26	2216	68.7	663	2	T38155
27	2214.5	68.7	608	2	T05741
28	2200	68.2	666	2	T06358
29	2173	67.4	682	1	H8BYK2

30	2156	66.9	663	2	S20877	dnak-type molecule
31	2108	65.4	651	2	JC7132	heat shock protein
32	2100	65.1	655	2	A48468	dnak-type molecule
33	2083	64.6	646	2	S07197	dnak-type molecule
34	2083	64.6	646	2	JC4853	dnak-type molecule
35	2083	64.6	646	2	A27077	dnak-type molecule
36	2080	64.5	641	2	PC7036	heat shock protein
37	2079	64.5	647	2	T41121	heat shock protein
38	2078	64.5	650	2	S11456	dnak-type molecule
39	2078	64.4	645	2	S41372	dnak-type molecule
40	2077	64.4	646	2	A45935	dnak-type molecule
41	2076.5	64.4	640	2	S37394	dnak-type molecule
42	2074.5	64.3	648	2	S53498	dnak-type molecule
43	2074.5	64.3	656	2	S51712	dnak-type molecule
44	2074	64.3	646	2	A35922	dnak-type molecule
45	2072	64.2	639	2	S20139	dnak-type molecule

## ALIGNMENTS

## RESULT 1

A27414  
dnak-type molecular chaperone GRP78 precursor - Chinese hamster  
N:Alternate names: glucose-regulated 78k protein; Ig heavy chain-binding protein  
C:Species: Cricetus griseus (Chinese hamster)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 20-Aug-1999  
C:Accession: A27414  
R:Ring, J.; Wooden, S.K.; Kriz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.  
Gene 55, 147-152, 1987  
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei  
A:Reference number: A27414; MUID:87305586  
A:Accession: A27414  
A:Molecule type: mRNA  
A:Residues: 1-654 <TIN>  
A:Cross-references: GB:M17169; NID:9191090; PID:AAA51448.1; PID:9304510  
C:Genetics:  
A:Gene: GRP78  
C:Function:  
A:Description: involved in protein folding and assembling/disassembling of protein co  
C:Superfamily: heat shock protein 70.  
C:Keywords: ATP; heterotrimer; immunoglobulin; molecular chaperone  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-654/Product: Ig heavy chain binding protein #status predicted <MAT>  
F:651-654/Region: endoplasmic reticulum retention signal

Query Match 99.5%; Score 3209; DB 2; Length 654;  
Best Local Similarity 99.7%; Pred. No. 4e-155;  
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	2	EDDKEDVGVVIGLGTTCVGVFKNGRVEILINDGSRITTPSYVATPGERLIGDA	61
DB	20	EDDKEDVGVVIGLGTTCVGVFKNGRVEILINDGSRITTPSYVATPGERLIGDA	79
QY	62	AKNQLTSPNENTVPAKRLIGRTWMDPSVQODIKELPFVVEKTKPYIQVDIGGGQRT	121
DB	80	AKNQLTSPNENTVPAKRLIGRTWMDPSVQODIKELPFVVEKTKPYIQVDIGGGQRT	139
QY	122	PAPEETISAVLTFRKETAAYLGKVTAVVTPAVFENDARQATKDACTAGLNMRII	181
DB	140	PAPEETISAVLTFRKETAAYLGKVTAVVTPAVFENDARQATKDACTAGLNMRII	199
QY	182	NEPTAAIAYGIDKRGSENIIVFDGGGTFPVSLITDNGVEVAVTNGDHLGGEDPD	241
DB	200	NEPTAAIAYGIDKRGSENIIVFDGGGTFPVSLITDNGVEVAVTNGDHLGGEDPD	259
QY	242	QRMEHFILYKKTKGVDKRNRAVQKLRREVEKAKRALSSOHQARIESEFSEGEDPS	301
DB	260	QRMEHFILYKKTKGVDKRNRAVQKLRREVEKAKRALSSOHQARIESEFSEGEDPS	319
QY	302	ETLTKRFEELMDLFRSTWKPQVKLESDIKSDIDEIVLVGGSTRIPKIQOLVKEFF	361

Dh 320 ETLTRAKFEELNMDLFRSTNKPQVKVLESDLKKSDIDEIVLVGSGTRIPKIQOLVKEFF 379  
Qy 362 NGKEPSRGINPDEAVNAGAAVQAGVLSGDDDTGDLVLLDVCPPLTGLGIEVGVWTKLIPR 421  
Dh 380 NGKEPSRGINPDEAVNAGAAVQAGVLSGDDDTGDLVLLDVCPPLTGLGIEVGVWTKLIPR 439  
Qy 422 NTVVPTKKSQIFSTASDNQPTVTIKVEGERPLTKDNHLLGTDFDLGIPAPRGVQIEV 481  
Dh 440 NTVVPTKKSQIFSTASDNQPTVTIKVEGERPLTKDNHLLGTDFDLGIPAPRGVQIEV 499  
Qy 482 TFEIDVNGILRVTAEDKGTGNKKKITTTNDQNLTPREIRMYNDAKFAEEDKKLKERI 541  
Dh 500 TFEIDVNGILRVTAEDKGTGNKKKITTTNDQNLTPREIRMYNDAKFAEEDKKLKERI 559  
Qy 542 DTRNELESAVSLKNOIGDKKLGKLSSEDKETMEKAVEKEIEMLESHODADIEDPKAK 601  
Dh 560 DTRNELESAVSLKNOIGDKKLGKLSSEDKETMEKAVEKEIEMLESHODADIEDPKAK 619  
Qy 602 KKELEEVOPRIISKLYGSAGPPPTGEEDTAE 632  
Dh 620 KKELEEVOPRIISKLYGSAGPPPTGEEDTSE 650

## RESULT 2

A37048  
dnaK-type molecular chaperone grp78 precursor - mouse  
N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein  
C:Species: Mus musculus (house mouse)  
C>Date: 15-Jan-1993 #sequence, revision 15-Jan-1993 #text, change 20-Aug-1999  
C:Accession: A37048; B37048; J00094; I49137; A31934  
R:Koutsumi, Y.; Norimington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.  
J. Cell Sci. Suppl. 11, 115-137, 1989  
A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulated  
A:Reference number: A37048; M0ID:90130686  
A:Accession: A37048  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-655 <K21>  
A:Accession: B37048  
A:Molecule type: protein  
A:Residues: 20-41, 'X', 43-46 <K22>  
R:Parfitt, C.L.J.; Holbuer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.  
Gene 82, 291-303, 1989  
A:Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero  
A:Reference number: J00094; M0ID:90060818  
A:Accession: J00094  
A:Molecule type: mRNA  
A:Residues: 488-590, 'G', 592-595, 'E', 597-655 <PAR>  
A:Cross-references: GB:M30779; NID:g193644; PIDN:AAA37742.1; PID:g193645  
R:Tillman, J.B.; Mote, P.L.; Walford, R.L.; Spindler, S.R.  
Gene 158, 225-229, 1995  
A:Title: Structure and regulation of the mouse Grp78 (B1P) promoter by glucose and calci  
A:Reference number: I49137; M0ID:95331621  
A:Accession: I49137  
A>Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-28 <RES>  
A:Cross-references: EMBL:U016277; NID:g829364; PIDN:AAA76734.1; PID:g829365  
R:Haas, I.G.; Meo, T.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988  
A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.  
A:Reference number: A31934; M0ID:8816922  
A:Accession: A31934  
A:Molecule type: mRNA  
A:Residues: 514-589, 'D', 591-595, 'E', 597-655 <HAAS>  
A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113  
C:Genetics:  
A:Gene: grp78  
C:Function:  
A:Description: Involved in protein folding and assembling/disassembling of protein comp  
C:Superfamily: heat shock protein 70  
C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon  
F:1-19/Domain: signal sequence #status predicted <Sig>

F:20-655/Product: dnaK-type molecular chaperone grp78 #status experimental <Mat>  
F:652-655/Region: endoplasmic reticulum retention signal

Query Match 99.3%; Score 3201; DB 2; Length 655;  
Best Local Similarity 99.4%; Pred. No. 1e-154;  
Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEDKKEDVGVTVGIDIGTTCVGVFKNGRVEILANDQGRIRTPSVAFPEGERLIGDA 61  
Dh 21 EEDKKEDVGVTVGIDIGTTCVGVFKNGRVEILANDQGRIRTPSVAFPEGERLIGDA 80  
Qy 62 AKNQLTSNPENTPFDKAKRLIGRTMNPVSVOODIKPLPFKVEKTKRYIYOVDIGGQTKT 121  
Dh 81 AKNQLTSNPENTPFDKAKRLIGRTMNPVSVOODIKPLPFKVEKTKRYIYOVDIGGQTKT 140  
Qy 122 FAPBEISAMVLTVMKETAETAEYLVCKKVTHAVVTPAYFNDQROATYDAGTIAGLNVRIT 181  
Dh 141 FAPBEISAMVLTVMKETAETAEYLVCKKVTHAVVTPAYFNDQROATYDAGTIAGLNVRIT 200  
Qy 182 NEPTAAAIAGYLRKREGENILVFDIGGTFEDVSLITDNGREVVATNGDTHLGEDFD 241  
Dh 201 NEPTAAAIAGYLRKREGENILVFDIGGTFEDVSLITDNGREVVATNGDTHLGEDFD 260  
Qy 242 QRYMEHFILKYYKKTKGKDVKKDNRAVQKLRREVEKAKRALSSQHOARIEIESFEDEFS 301  
Dh 261 QRYMEHFILKYYKKTKGKDVKKDNRAVQKLRREVEKAKRALSSQHOARIEIESFEDEFS 320  
Qy 302 ETLTRAKFEELNMDLFRSTNKPQVKVLESDLKKSDIDEIVLVGSGTRIPKIQOLVKEFF 361  
Dh 321 ETLTRAKFEELNMDLFRSTNKPQVKVLESDLKKSDIDEIVLVGSGTRIPKIQOLVKEFF 380  
Qy 362 NGKEPSRGINPDEAVNAGAAVQAGVLSGDDDTGDLVLLDVCPPLTGLGIEVGVWTKLIPR 421  
Dh 381 NGKEPSRGINPDEAVNAGAAVQAGVLSGDDDTGDLVLLDVCPPLTGLGIEVGVWTKLIPR 440  
Qy 422 NTVVPTKKSQIFSTASDNQPTVTIKVEGERPLTKDNHLLGTDFDLGIPAPRGVQIEV 481  
Dh 441 NTVVPTKKSQIFSTASDNQPTVTIKVEGERPLTKDNHLLGTDFDLGIPAPRGVQIEV 500  
Qy 482 TFEIDVNGILRVTAEDKGTGNKKKITTTNDQNLTPREIRMYNDAKFAEEDKKLKERI 541  
Dh 501 TFEIDVNGILRVTAEDKGTGNKKKITTTNDQNLTPREIRMYNDAKFAEEDKKLKERI 560  
Qy 542 DTRNELESAVSLKNOIGDKKLGKLSSEDKETMEKAVEKEIEMLESHODADIEDPKAK 601  
Dh 561 DTRNELESAVSLKNOIGDKKLGKLSSEDKETMEKAVEKEIEMLESHODADIEDPKAK 620  
Qy 602 KKELEEVOPRIISKLYGSAGPPPTGEEDTAE 632  
Dh 621 KKELEEVOPRIISKLYGSAGPPPTGEEDTSE 651

## RESULT 3

HHR7GB  
dnaK-type molecular chaperone precursor - rat  
N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein  
N:Contains: steroidogenesis-activator polypeptide  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Mar-1993 #sequence, revision 31-Mar-1993 #text, change 22-Jun-1999  
C:Accession: A23948; A60134; A26257  
R:Munro, S.; Pelham, H.R.B.  
Cell 46, 291-300, 1986  
A:Title: An Hsp70-like protein in the ER: identitly with the 78 kd glucose-regulated p  
A:Reference number: A23948; M0ID:8625075  
A:Accession: A23948  
A:Molecule type: mRNA  
A:Residues: 1-654 <MUN>  
A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA0817.1; PID:g203151  
R:Pedersen, R.C.; Browne, A.C.  
Science 226, 188-190, 1987  
A:Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor  
A:Reference number: A60134; M0ID:87177981

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 23.81 Seconds  
(without alignments)  
603.931 Million cell updates/sec

Title: US-09-806-955-1

Perfect score: 3273

Sequence: 1 MEEDKREKDVGVWIDLGTT.....AGPPTGEEDTAELHHHHH 639

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	98.0	654	1 US-08-441-139-11	Sequence 11, Appl
2	3168	96.8	666	1 US-08-441-139-16	Sequence 16, Appl
3	2174	66.4	682	1 US-08-441-139-2	Sequence 2, Appl
4	2156	65.9	663	1 US-08-441-139-7	Sequence 7, Appl
5	2077	63.5	646	1 US-08-441-139-14	Sequence 14, Appl
6	2065.5	63.1	679	1 US-08-441-139-5	Sequence 5, Appl
7	1938	59.2	643	4 US-08-797-358B-3	Sequence 3, Appl
8	1547.5	47.3	679	1 US-08-214-583-2	Sequence 2, Appl
9	1474	45.0	641	1 US-08-441-139-4	Sequence 4, Appl
10	1461.5	44.7	649	4 US-09-066-047-5	Sequence 5, Appl
11	1444	44.1	607	2 US-08-472-534-5	Sequence 5, Appl
12	1303.5	39.8	600	6 5240706-1	Patent No. 5240706
13	1140.5	34.8	339	2 US-08-928-692-52	Sequence 52, Appl
14	941	28.8	187	6 5196523-13	Patent No. 5196523
15	824	25.2	168	1 US-08-441-139-10	Sequence 10, Appl
16	818.5	25.0	315	1 US-08-257-073-7	Sequence 7, Appl
17	750.5	22.9	471	1 US-08-203-905B-2	Sequence 2, Appl
18	726.5	22.2	472	1 US-08-203-905B-14	Sequence 14, Appl
19	642.5	19.6	139	6 5196523-10	Patent No. 5196523
20	607.5	18.6	999	2 US-08-770-301A-3	Sequence 3, Appl
21	607.5	18.6	999	3 US-09-175-581-3	Sequence 3, Appl
22	598	18.3	999	2 US-08-770-301A-1	Sequence 1, Appl
23	528.5	16.1	999	3 US-09-175-581-1	Sequence 53, Appl
24	437.5	13.4	374	2 US-08-928-692-51	Sequence 51, Appl
25	402	12.3	77	6 5196523-7	Patent No. 5196523
26	395	12.1	79	6 5196523-11	Patent No. 5196523

28	367.5	11.2	599	4 US-09-080-983-9	Sequence 9, Appl
29	358	10.9	549	2 US-08-770-544-6	Sequence 6, Appl
30	345.5	10.6	80	1 US-08-464-164-4	Sequence 4, Appl
31	345.5	10.6	31	1 US-08-338-057-4	Sequence 4, Appl
32	345.5	10.6	80	2 US-08-668-416-4	Sequence 4, Appl
33	323	9.9	554	5 PCT-US94-06430-7	Sequence 7, Appl
34	320	9.8	554	4 US-08-591-468-7	Sequence 7, Appl
35	267.5	8.2	714	2 US-08-472-534-3	Sequence 3, Appl
36	264	8.1	56	6 5196523-12	Patent No. 5196523
37	230	7.0	46	6 5196523-8	Patent No. 5196523
38	211	6.4	42	6 5196523-15	Patent No. 5196523
39	192	5.9	38	6 5196523-9	Patent No. 5196523
40	168.5	5.1	1786	4 US-08-973-462-8	Sequence 8, Appl
41	151.5	4.6	1805	1 US-07-853-913-2	Sequence 2, Appl
42	147	4.5	693	4 US-08-235-836C-68	Sequence 68, Appl
43	142.5	4.4	2285	4 US-09-308-375-2	Sequence 2, Appl
44	138	4.2	700	1 US-07-720-589-2	Sequence 2, Appl
45	138	4.2	700	2 US-08-785-190-2	Sequence 2, Appl

# ALIGNMENTS

```

RESULT 1
US-08-441-139-11
; Sequence 11, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: WILTRUP, Dr. Karl D.
; APPLICANT: ROBINSON, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,139
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-139-11

Query Match 98.0%; Score 3209; DB 1; Length 654;
Best Local Similarity 99.7%; Pred. No. 6.1e-250;
Matches 629; Conservative 0; Indels 0; Gaps 0;

```

QY	2	EDKKEDVGTVAAGDIDGTTTSCVCFVKNGKRELIANOGNRTPTSYAALFPEGGRLLGDA	61
Db	20	EDKKEDVGTVAAGDIDGTTTSCVCFVKNGKRELIANOGNRTPTSYAALFPEGGRLLGDA	79
QY	62	AKNOLTSNPENTVFDAKRLIGRTWNDSVQODIKFLPFKVVEKTKPYIOVDIGGQTKT	121
Db	80	AKNOLTSNPENTVFDAKRLIGRTWNDSVQODIKFLPFKVVEKTKPYIOVDIGGQTKT	139
QY	122	FAPEISAMVLTKKKETAALGKRVHAVVTVAYATVENDAQROTKACGTAGLNVARI	181
Db	140	FAPEISAMVLTKKKETAALGKRVHAVVTVAYATVENDAQROTKACGTAGLNVARI	199
QY	182	NEPTAAALAYGLDKREGEKNILTVEDLGCGTFDYSLLTIDNGVEFWATNGDTHLGSGDFD	241
Db	200	NEPTAAALAYGLDKREGEKNILTVEDLGCGTFDYSLLTIDNGVEFWATNGDTHLGSGDFD	258
QY	242	QVMEHFITKLXKKTKGDKVRKDNRAVOKLRREVEKAKRALSSQOARIEIESFEYGEDFS	301
Db	260	QVMEHFITKLXKKTKGDKVRKDNRAVOKLRREVEKAKRALSSQOARIEIESFEYGEDFS	319
QY	302	ETLFRAKREELNMLFRSTMPQVQVLEDSLKKSDIDEIYVCGSTRIRKIQQLVKEFF	366
Db	320	ETLFRAKREELNMLFRSTMPQVQVLEDSLKKSDIDEIYVCGSTRIRKIQQLVKEFF	379
QY	362	NGKESRGINDEDEAAVAAVQAGVLSGDDQDTGDLVLLDVCPLTGLTETGVGWTKLIPR	421
Db	380	NGKESRGINDEDEAAVAAVQAGVLSGDDQDTGDLVLLDVCPLTGLTETGVGWTKLIPR	438
QY	422	NTVVFTRKSKQIFSTASDNQPTVTITKYVEGERPLTKDNHLLGTFDLTGIPAPRGVQIEV	481
Db	440	NTVVFTRKSKQIFSTASDNQPTVTITKYVEGERPLTKDNHLLGTFDLTGIPAPRGVQIEV	499
QY	482	TFEIDVNGILVATFADKGTGKKNKITTINDQNRLLTPEIEIKMNVDAKFAEDKKLKERI	544
Db	500	TFEIDVNGILVATFADKGTGKKNKITTINDQNRLLTPEIEIKMNVDAKFAEDKKLKERI	558
QY	542	DTNRELSEYASLKNQIQIDKKEKLGKLSSEDEKTEMEKAVEEIKMLSHODADIEDPKAK	601
Db	560	DTNRELSEYASLKNQIQIDKKEKLGKLSSEDEKTEMEKAVEEIKMLSHODADIEDPKAK	619
QY	602	KKELEIVQPIISKLYGSAGPPPTGEEEDTAE	632
Db	620	KKELEIVQPIISKLYGSAGPPPTGEEEDTSE	650
RESULT 2			
US-08-441-139-16			
Sequence 16, Application US/08441139			
Patent No. 5773245			
GENERAL INFORMATION:			
APPLICANT: Wiltup, Dr. Karl D.			
APPLICANT: Robinson, Anne S.			
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF			
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS			
NUMBER OF SEQUENCES: 20			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER			
STREET: 400 Garden City Plaza			
CITY: Garden City			
STATE: NY			
COUNTRY: USA			
ZIP: 11530			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/441,139			
FILING DATE: 15-MAY-1995			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			

```

? APPLICATION NUMBER: US 08/089, 997
? FILING DATE: 06-JUL-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Digillo, Frank S.
? REGISTRATION NUMBER: 31,346
? REFERENCE/DOCKET NUMBER: 8646
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 516-742-4343
? TELEFAX: 516-742-4366
?
? TELE: 230 901 SANS UR
? INFORMATION FOR SEQ ID NO: 16:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 666 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
? MOLECULE TYPE: protein
?
? US-08-441-139-16

```

Query Match	Similarity	97.8%	Score	1168:	DB	1:	Length	666:
Best Local	Similarity	97.8%	Pred.	Mis.	1.3e-245:			
Matches	617:	Conservative	11:	No. Matches	5:	Indels	0:	Gaps
QY	2	EEDEKEDVTVVGGIDLGTTYSCVGVFKNKRVEIIANDQGNRTTPSVAFTEGGERLIGDA	61					
DB	32	DEEKEDVGTGVGIDLGTTYSCVGVFKNKRVEIANDQGNRTTPSVAFTEGGERLIGDA	91					
QY	62	AKNQLTSPNEMTVFPAKRLIGRTWMDPSVQODIKFLPEKVEVKTKPTVQVDIGGQRT	121					
DB	92	ANNQLTSPNEMTVFPAKRLIGRTWMDPSVQODIKFLPEKVEVKTKPTVQVVGQQT	151					
QY	122	FAPDEISAAVNLTKMKEETAAYLGKKVYTAHVYVVPAYFNDAQARQADACTIGLAVNR	181					
DB	152	FAPDEISAAVNLTKMKEETAAYLGKKVYTAHVYVVPAYFNDAQARQADACTIGLAVNR	211					
QY	182	NEPTAAAIYAGLDKREGEKNILVFDLGCGTDPVSLTTIDNGVEEVVATNGDTHLGSEDD	241					
DB	212	NEPTAAAIYAGLDKREGEKNILVFDLGCGTDPVSLTTIDNGVEEVVATNGDTHLGSEDD	271					
QY	242	QVMEHFITLYKKTKGKDYRKDNRAVQKLREVEKAKRALSSQHOARIEISFEGEDS	301					
DB	272	QVMEHFITLYKKTKGKDYRKDNRAVQKLREVEKAKRALSSQHOARIEISFEGEDS	331					
QY	302	EFLTRAKFEELWMDLFRSTMKPVOKVLESDLSKSDIDETIVYGGSTRIPKIQOLYKEF	361					
DB	332	EFLTRAKFEELWMDLFRSTMKPVOKVLESDLSKSDIDETIVYGGSTRIPKIQOLYKEF	391					
QY	362	NCKEFSRGINPDEAAVAYGAAGVLSGDDQDGLDLVDVCEPLTLGIEYGVGVMTKLIPR	421					
DB	392	NCKEFSRGINPDEAAVAYGAAGVLSGDDQDGLDLVDVCEPLTLGIEYGVGVMTKLIPR	451					
QY	422	NTVPTPKSQIFSTASDNOPIVTIKYVBSERPLTKDNHLLGTFDLGTIPAPRGVQIIV	481					
DB	452	NTVPTPKSQIFSTASDNOPIVTIKYVBSERPLTKDNHLLGTFDLGTIPAPRGVQIIV	511					
QY	482	TPEIDVNGILRTAEDKCGKNKTKTTITNDONRLPPEELERNVANDKKAABEDKKLKEKI	541					
DB	512	TPEIDVNGILRTAEDKCGKNKTKTTITNDONRLPPEELERNVANDKKAABEDKKLKEKI	571					
QY	542	DPNRELSTAYSLKNOIGDKERLGGKLSSEDEETMEKAAVEERIEWLESHQDADIEDFKAK	601					
DB	572	DARNELESAYSLKNOIGDKERLGGKLSSEDEETIEKAAVEERIEWLESHQDADIEDFKAK	631					
QY	602	KKELEIVQPIISKLYGSAGPPTGGEEDTAE	632					
DB	632	KKELEIVQPIISKLYGSAGPPTGGEEDTAE	662					

<u>DB Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
USPT	recombinant adj immunoglobulin adj heavy adj chain adj binding adj protein	0	<u>L12</u>
USPT	recombinant adj (Ig) adj heavy adj chain adj binding adj protein	0	<u>L11</u>
USPT,PGPB	13 and (pharmaceutic\$ adj composition\$)	15	<u>L10</u>
USPT,PGPB	16 and pharmaceutic\$	39	<u>L9</u>
USPT,PGPB	13 and inflamm\$	14	<u>L8</u>
USPT,PGPB	13 and pharmaceutic\$	46	<u>L7</u>
USPT,PGPB	13 and composition\$	72	<u>L6</u>
USPT,PGPB	14 same composition\$	0	<u>L5</u>
USPT,PGPB	12 same recombinant\$	4	<u>L4</u>
USPT,PGPB	12 and recombinant\$	94	<u>L3</u>
USPT,PGPB	((Ig) or immunoglob\$) adj heavy adj chain adj binding adj protein\$	113	<u>L2</u>
USPT,PGPB	recombinant adj heavy adj chain adj binding adj protein\$	0	<u>L1</u>

**THIS PAGE BLANK (USPTO)**



(FILE 'HOME' ENTERED AT 10:33:09 ON 07 DEC 2001)

FILE 'BIOSIS, MEDLINE, EMBASE, CAPLUS' ENTERED AT 10:33:22 ON 07 DEC 2001  
L1 784 S IMMUNOGLOBULIN? (W) HEAVY (W) CHAIN (W) BINDING (W) PROTEIN?  
L2 20 S L1 (P) RECOMBINANT  
L3 8 DUP REM L2 (12 DUPLICATES REMOVED)

=>

09/806,955

**THIS PAGE BLANK (USPTO)**

L2 ANSWER 1 OF 20 BIOSIS COPYRIGHT 2001 BIOSIS  
TI Effects of co-expressing chaperone BiP on functional antibody production  
in the baculovirus system.  
AU Hsu, Tsu-An; Eiden, Joseph J.; Bourgarel, Pierre; Meo, Tommaso;  
Betenbaugh, Michael J. (1)  
SO Protein Expression and Purification, (1994) Vol. 5, No. 6, pp. 595-603.  
ISSN: 1046-5928.  
AB The assembly pathway of the insect cell Spodoptera frugiperda (Sf-9) was  
engineered to include expression of the murine chaperone  
**immunoglobulin heavy chain binding**  
**protein** (BiP) using the baculovirus vector. The impact of BiP  
coexpression on the production and secretion of functional and soluble  
**recombinant** immunoglobulin IgG levels was evaluated.  
**Recombinant** BiP was found to associate specifically with  
immunoglobulins in immunoprecipitation studies. Coinfection of insect  
cells with a BiP-containing baculovirus and baculoviruses coding for two  
different murine IgG proteins increased intracellular functional antibody  
activity levels substantially above the levels observed in the absence of  
BiP. Soluble intracellular immunoglobulin levels were found to increase as  
well. However, secreted functional antibody levels did not increase  
significantly. Also, degradation of heavy chain immunoglobulin in insect  
cells was indicated by the accumulation of lower molecular weight  
immunoglobulins at 4 days postinfection. Coexpression of light chains  
reduced the level of these lower molecular weight immunoglobulins while  
BiP coexpression led to enhanced levels. These findings suggest that  
coexpressed BiP can increase intracellular soluble and functional antibody  
yields but that secretion in the baculovirus-insect cell system must be  
limited at some post-translational step.

**THIS PAGE BLANK (USPTO)**

L2 ANSWER 18 OF 20 CAPLUS COPYRIGHT 2001 ACS  
TI Modification of hepatic immunoglobulin heavy chain binding protein  
(BiP/Grp78) following exposure to structurally diverse peroxisome  
proliferators  
AU Witzmann, F. A.; Jarnot, B. M.; Parker, D. N.; Clack, J. W.  
SO Fundam. Appl. Toxicol. (1994), 23(1), 1-8  
CODEN: FAATDF; ISSN: 0272-0590  
AB This investigation was conducted to det. the comparative effect of  
structurally diverse peroxisome proliferators (PP) on the two-dimensional  
protein pattern of rat liver whole homogenates. Perfluoro-n-decanoic acid  
(PFDA), perfluoro-n-octanoic acid (PFOA), clofibrate, and di(2-ethylhexyl)  
phthalate (DEHP) are all known to cause the proliferation of hepatic  
peroxisomes and the induction of peroxisomal .beta.-oxidative and  
microsomal .omega.-oxidative enzymes. To clarify the mechanistic  
differences between these compds. with regard to the liver, the authors  
examd. the unique patterns of protein alteration produced by in vivo  
exposure to them. Following exposure to various doses, whole liver  
homogenates were prepd. and sepd. by two-dimensional gel electrophoresis  
(2DE) using to the ISO-DALT system. Stained gels were digitized and  
protein patterns analyzed using the Kepler 2D gel anal. system.  
**Ig heavy chain binding**  
**protein** (BiP), also known as 78-kDa glucose-regulated protein  
(Grp78), was identified immunol. and by comigration of **recombinant**  
Grp78. BiP is a luminal endoplasmic reticular protein that functions is  
the assembly and folding of nascent proteins as they enter the ER. The  
present results suggest a selective posttranslational modification of BiP  
following PFDA exposure. Single-dose exposure to PFDA was assocd. with a  
notable charge modification of BiP that persists up to 30 days. PFOA,  
clofibrate, and DEHP had less effect in this regard. The identity of  
BiP/Grp78 as the halothane hepatitis-assocd. trifluoroacetylated protein  
was also demonstrated. The nature of this PFDA-assocd. protein  
modification (reactive metabolite conjugation, abnormal ribosylation, or  
phosphorylation) is currently under investigation. The results document  
PFDA's unique toxicity as a PP and support the utility of 2D gel anal. in  
toxicity testing.

**THIS PAGE BLANK (USPTO)**

- TI cDNA cloning of the immunoglobulin heavy chain binding protein  
AU Haas, Ingrid G.; Meo, Tommaso  
SO Proc. Natl. Acad. Sci. U. S. A. (1988), 85(7), 2250-4  
CODEN: PNASA6; ISSN: 0027-8424
- AB A cDNA library was constructed from size-fractionated poly(A)+ RNA prep'd. from a murine pre-B-cell hybridoma expressing high levels of **Ig heavy-chain-binding protein (BiP)** and  $\mu$  heavy chains. Transformed bacterial colonies were screened for **recombinant** plasmids contg. cDNA coding for BiP by hybrid-selected mRNA translation. A clone, pMBiP, contg. a 736-base-pair insert was shown to encode the protein. Translation in vitro of hybridoma mRNA selected by hybridization to the pMBiP cDNA yielded a single polypeptide of BiP-like size. The authenticity of this mRNA was verified by comparing the peptides obtained by the limited proteolysis of its in vitro translation product with those obtained from the in vivo produced BiP. Likewise, the authenticity of the cDNA insert was verified by an RNase A protection assay of heteroduplex mols. obtained by annealing a uniformly labeled single-strand copy of the cDNA clone with the same mRNA selected by hybridization and tested by translation. The nucleotide sequence of this clone was used to deduce the C-terminal 142 amino acids of BiP and to establish its kinship with the 70-kDa heat-shock protein family. The finding of a single copy of the BiP gene in DNA blots of mouse and rat implies that the BiP-related RNA transcripts constitutively expressed in various murine tissues and cell lines are indeed products of the same gene. These findings imply that BiP plays a more general role than previously anticipated on the basis of the discovery of its assocn. with Ig heavy chains.

**THIS PAGE BLANK (USPTO)**



L2 ANSWER 8 OF 20 MEDLINE

TI Identification of immunoglobulin heavy chain binding protein as glucose-regulated protein 78 on the basis of amino acid sequence, immunological cross-reactivity, and functional activity.

AU Kozutsumi Y; Normington K; Press E; Slaughter C; Sambrook J; Gething M J  
SO JOURNAL OF CELL SCIENCE. SUPPLEMENT, (1989) 11 115-37.

Journal code: HNG; 8502898. ISSN: 0269-3518.

AB **Immunoglobulin heavy chain binding**

**protein** (BiP) associates transiently with various proteins destined for the secretory pathway. To investigate the relationship between BiP and the 78K (K = 10(3) Mr) glucose-regulated protein (GRP78), we have determined a partial amino acid sequence of purified mouse BiP and isolated and sequenced a full-length cDNA clone encoding mouse GRP78. The 26 amino-terminal residues of the mature BiP protein are identical to a sequence of amino acids located near the start of the open reading frame encoding GRP78. A polyclonal antiserum raised against mouse GRP78 protein expressed in bacteria from the cloned GRP78 cDNA could immunoprecipitate complexes consisting of BiP and unfolded forms of immunoglobulin heavy chains. Furthermore, a monoclonal antibody raised against mouse BiP immunoprecipitated mouse GRP78 expressed in monkey CV-1 cells from an SV40-GRP78 **recombinant** vector. Finally, like the endogenous BiP of simian cells, mouse GRP78 associated with malformed, non-glycosylated forms of influenza hemagglutinin (HA) when GRP78 and HA were co-expressed from SV40 vectors in CV-1 cells. These studies confirm that BiP is identical to GRP78. Comparison of the nucleic acid and deduced amino acid sequence of mouse GRP78 with those of other rodent and human GRP78s revealed an extremely high degree of sequence identity. BiP/GRP78 is closely related (approximately 60% identity) to the cytoplasmic 70K heat-shock proteins. Surprisingly, the carboxy-terminal 29 amino acids of BiP/GRP78, which are not conserved in HSP70 proteins, are almost identical in sequence to the steroidogenesis activator peptide found in the cytoplasm of rat Leydig tumor cells. Possible relationships between these polypeptides are discussed.

**THIS PAGE BLANK (USPTO)**